



# On the working memory of humans and great apes: Strikingly similar or remarkably different?

Dwight W. Read<sup>a,b,\*</sup>, Héctor M. Manrique<sup>c</sup>, Michael J. Walker<sup>d</sup>

<sup>a</sup> Department of Anthropology, University of California Los Angeles, Los Angeles, CA, 90095, USA

<sup>b</sup> Department of Statistics, University of California Los Angeles, Los Angeles, CA, 90095, USA

<sup>c</sup> Departamento de Psicología y Sociología, Universidad de Zaragoza, Zaragoza, Spain

<sup>d</sup> Department of Zoology and Physical Anthropology, Faculty of Biology, University of Murcia, Murcia, Spain

## ARTICLE INFO

### Keywords:

Working memory  
Human evolution  
Cognitive evolution  
Comparative psychology  
Chimpanzee  
Hominin evolution  
Theory of mind  
Planning

## ABSTRACT

In this article we review publications relevant to addressing widely reported claims in both the academic and popular press that chimpanzees working memory (WM) is comparable to, if not exceeding, that of humans. WM is a complex multidimensional construct with strong parallels in humans to prefrontal cortex and cognitive development. These parallels occur in chimpanzees, but to a lesser degree. We review empirical evidence and conclude that the size of WM in chimpanzees is  $2 \pm 1$  versus Miller's famous  $7 \pm 2$  in humans. Comparable differences occur in experiments on chimpanzees relating to strategic and attentional WM subsystems. Regardless of the domain, chimpanzee WM performance is comparable to that of humans around the age of 4 or 5. Next, we review evidence showing parallels among the evolution of WM capacity in hominins ancestral to *Homo sapiens*, the phylogenetic evolution of hominins leading to *Homo sapiens*, and evolution in the complexity of stone tool technology over this time period.

## 1. Introduction

The last decade has seen a dramatic change in the way we think about commonalities between us and the pongids. This change in thinking has been driven by experimental discoveries that attribute to them many cognitive traits traditionally reserved for humans. It has been reported that pongids can save and make tools for future use (Bräuer and Call, 2015; Mulcahy and Call, 2006), are able to adopt another individual's visual perspective (Hare et al., 2000; Karg et al., 2015), attribute knowledge of the world to conspecifics (Kaminski et al., 2008), and perhaps attribute beliefs to humans (Krupenye et al., 2016). They also have been observed to cooperate to solve common feeding problems, both in the wild (Boesch, 1994, 2002) and in captivity (Melis et al., 2006, 2011). Moreover, complex planning and the capacity to remember past personal events are claimed to be among their cognitive repertoire (Lewis et al., 2019; Mulcahy and Call, 2006; Martin-Ordas et al., 2010). There are even studies reporting that they may incur personal costs to witnessing how justice is dispensed (Mendes et al., 2018). When perusing these studies, one is reminded of Louis Leakey's comment when Jane Goodall informed him that chimpanzees used and made tools, "Now we must redefine tool, redefine Man, or accept

chimpanzees as humans" (Goodall, 1998). If only using tools made Leakey question whether chimpanzees belong to our own species, what would he say now that many cognitive frontiers thought to separate chimpanzees from humans have been trespassed? Would he be forced to admit that chimpanzees are indeed humans? This *reductio ad absurdum* begs as many questions as it purports to answer. How might we resolve the matter?

At the neurological level, comparable brain structures in humans, chimpanzees, and other pongids, suggest evolutionary continuity, though that by no means excludes the possibility of qualitative differences between them (cf., Buxhoeveden and Casanova, 2002; Preuss, 2004; Rilling, 2006; Rilling et al., 2007; Semendeferi et al., 2001, among others). Even between the closely related chimpanzee (*Pan troglodytes*) and bonobo (*Pan paniscus*), small differences in their respective neurological systems relating to empathy and aggression may underlie striking differences in their social behavior (Rilling et al., 2012). Thus, while there is commonality in psychoneural processes in human and non-human primates (and other mammals), this need not translate into equivalent commonality for cognitive capacities and social behavior (Read, 2012a). In the last several decades, comparative psychologists have devoted a great deal of energy and resources to investigating

\* Corresponding author.

E-mail address: [dread@ss.ucla.edu](mailto:dread@ss.ucla.edu) (D.W. Read).

<https://doi.org/10.1016/j.neubiorev.2021.12.019>

Received 22 June 2020; Received in revised form 8 November 2021; Accepted 9 December 2021

Available online 14 December 2021

0149-7634/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

whether specific cognitive traits, once thought to be uniquely human, are also present in our closest living biological relatives. This approach has proved to be very productive for identifying commonalities between human and ape cognitive traits yet attempts at finding or explaining differences have not received the same attention. Here, we focus on working memory (WM) and its implications for cognitive performance and in so doing we are equally intent in explaining both what we (humans and apes) have in common and what might set us apart at the cognitive/behavioral level. We take a multidisciplinary approach in which findings from experiments relating to WM and conducted with captive apes and from implements made by archaic humans are reviewed, discussed, and interpreted through consideration of recent neuroscientific and archaeological developments, as well as detailed statistical comparisons.

It has been proposed that increase in the size of the holding capacity of working memory (HWM) makes possible those complex human cognitive abilities that allowed us to build our extremely complex technical and social worlds (Coolidge and Wynn, 2001; Read and van der Leeuw, 2008; Read et al., 2009; van der Leeuw, 2020; van der Leeuw et al., 2009; Wynn and Coolidge, 2004). If so, qualitative differences in cognitive abilities related to working memory capacity, including recursive reasoning (Read, 2006, 2008; cf., Corballis, 2011), could underlie the differences in the cognitive abilities and mental capacity of humans versus pongids, in general, and between humans and chimpanzees, in particular. We shall argue that this is the case and then review and evaluate the experimental evidence supporting our claim.

## 2. Working memory as a complex multidimensional concept

Working memory is a multi-faceted construct, the object of attention from psychologists and neuroscientists (Chai et al., 2018). Different theoretical models and their several components allegedly capture the essence of this construct and are implemented by observing behavior in response to recall, reasoning, planning, and/or problem-solving.

One recent account of working memory, drawing upon research that increasingly focuses on the quality and not just the quantity of what is kept active in WM, argues that WM should be regarded as being a “limited resource ... distributed flexibly among all items to be maintained in memory.... the quality rather than the quantity of working memory representations determines performance” (Ma et al., 2014:347). Of concern to Ma et al. are research results that focus in more detail on what is involved in memory recall and whether it is an entire entity or part of an entity that is the unit of recall and, if the latter, how these might relate to different kinds of resources. In brief, they consider that the shift from a focus on quantity to a focus on quality will “lead to a deeper understanding of how and why individuals remember and forget” (p. 355).

Consistent with this viewpoint is the idea of subsuming working memory in a neural Turing machine (cf., Turing, 1936), widely distributed in the brain, capable of carrying out a vast array of computations (Christophel et al., 2017; cf., Fuster, 2001, 2015). This would permit a wide range of items of information to be held at variable strengths and at variable levels of accessibility in working memory. From an information theory perspective, a characteristic of our cerebral neural components is their capacity for attaining a remarkable flexibility of bandwidth (Miller and Buschman, 2015), thus “information is encoded into working memory by allocating attention to internal representations” (D’Esposito and Postle, 2015:115). This connects with Baddeley’s (1996) view of the central executive component of working memory acting as a control system regarding selecting and rejecting streams of information for being acted on by working memory. In this regard, D’Esposito and Postle (2015) aver that Baddeley’s control system

... was not in any sense ‘specialized for’ or ‘dedicated to’ working memory operations, but one that could use and/or manipulate the

contents of working memory storage to more effectively guide behavior. The prefrontal, basal ganglia, thalamic, and brainstem systems ... can be construed as a neural substrate for this Central Executive (D’Esposito and Postle, 2015: 133).

They hold that confusion (from the 1980s to the 2000s) about whether the prefrontal cortex (PFC) acts as a buffer for WM or directly relates to the functioning of the central executive component of WM stems from widespread “misattribution of PFC activity to the functioning of one of the storage buffers from the multicomponent model rather than to the Central Executive” (D’Esposito and Postle, 2015: 133). Passingham and Wise (2012: 3): assert that “executive function ... has little to offer in the way of testable hypotheses,” that WM is a function of several regions in mammalian brains, and is demonstrably present in orders of mammals lacking prefrontal cortex (Passington and Wise, 2012:2–3, 312–315). Nor can WM be regarded as being the primary function of prefrontal cortex in primates (Ken-Ichiro et al., 2016).

However, determining the precise relationship of the PFC to working memory is not necessary for our goal of reviewing studies of the behavior of great apes, and relating their behavior to their working memory and how this relates to working memory in humans. For this reason, we do not bind ourselves to any specific theoretical model or author. Instead, we take to heart an article by Philip Beaman (2010) about the need to identify the sub-systems most often ascribed to working memory and their specific functions.

Beaman discusses three ways whereby working memory has been characterized. For our purposes here, these three main sub-systems contributing to working memory will provide the context for our discussion regarding the differences and similarities between humans and other non-human primates regarding WM. The three sub-systems identified by Beaman are as follows.

- 1 A holding sub-system for working memory (HWM) of sensory buffers for temporarily holding or storing rapidly decaying information as part of the manipulation of that information. The holding sub-system, whether it involves the buffer portion of the phonological loop, the visuospatial, or some other subsystem of working memory, holds activated information, some subset of which will be subjected to attentional control and processing by the executive function of working memory (Cowan, 1999; Engle et al., 1999). The holding sub-system has sometimes been referred to as short term memory but differs from short term memory in that it involves not only storing information temporarily but also acting upon this information (Diamond, 2013).

In the original Baddeley model for WM (Baddeley, 1986) there is a *spatial buffer* and a *phonological buffer*. More recently, Baddeley (2000) introduced an *episodic buffer* into his WM model when he recognized that his WM model did not include an explicit component for storage capacity (Baddeley, 2012). As he phrases it, the *episodic buffer* can “hold integrated episodes or chunks in multidimensional code” and “acts as a buffer store” (Baddeley, 2012: 15). For our purposes here, we incorporate his idea of buffers by referring, more generally, to a holding sub-system for working memory.

- 2 A strategic sub-system of working memory (SWM) for enabling the establishment of goals and sub-goals for guidance of behavior that involves “hierarchical processes controlling the order in which a sequence of operations is performed (Miller et al., 1960:16)” (Beaman, 2010: S27).
- 3 A working memory attentional sub-system (AWM) that allows focusing of attention onto significant features of one’s surroundings (Cowan, 1988, 1995, 1999, 2008; D’Esposito and Postle, 2015; Morey and Cowan, 2018; Rhodes and Cowan, 2018) as well as filtering out distractors. According to Beaman (2010: S28):

... the capability [of the AWM] to focus attention on task-relevant material and prohibit the processing of task-irrelevant material underpins performance on many cognitive tasks, including tests of reasoning and fluid intelligence (Conway et al., 2003).

In addition, he notes (p. S28) that the AWM relates to quantitative measures of WM:

Cowan (2001, 2005) provides an extended argument that only four items at once can be held within the focus of attention (Miller's [1956] earlier and larger estimate of  $7 \pm 2^1$  requires the utilization of more, and more variable, mnemonic strategies; Cowan et al., 2007). Cowan also considers this value to be relatively invariant ( $\pm 1$ ) within *Homo sapiens*.

Of these three sub-systems, the SWM and the AWM sub-systems are jointly roughly equivalent to Baddeley's central executive notion. By regarding the HWM sub-system to be a generalized version of Baddeley's episodic buffer, we can pay less attention to the ephemeral character of the information stored temporarily in the HWM sub-system than to the limited capacity of the HWM sub-system to hold the items of information upon which the ATM has focused attention, whether that capacity be as many as Miller's magical number 7, Cowan's 4, or even fewer, perhaps within a hierarchy of sub-systems (Cowan, 1988, 1995, 1999, 2001, 2005, 2008; McElree, 2006; Majerus et al., 2016, 2017; Miller, 1956; Morey and Cowan, 2018; Oberauer, 2013, 2018; Rhodes and Cowan, 2018).

We have, then, parsed WM into three separate sub-systems that facilitate our making fine-grain analyses and interpretations of the behavior of great apes undergoing experiments involving challenging cognitive tasks that sheds light on their WM abilities – or limitations – without necessarily reducing WM just to quantitative measures. Suffice it to say that we find the parts played in working memory by HWM, SWM and AWM are useful constructs for studying cognitively grounded behavior in the great apes. For this purpose, we have focused on experiments that provide accounts of behavior that can be related to these aspects of working memory.

The experiments that we have reviewed are shown in Table 1 in the order in which they are discussed. The bulk of our selections are studies that have investigated the cognition of chimpanzees in a way that is susceptible to being analyzed in terms of WM computational demands. In this regard, we selected all experiments that provide direct measurements of the size of HWM in chimpanzees. The latter is often the focus of comparisons between human and chimpanzee WM. In our selection, we included both natural experiments such as the chimpanzee nut-cracking at Bossou in Guinea, and experiments with monkeys when the latter relates to properties of WM in the great apes. We also included experiments that have been cited by authors in support of claims for extraordinary WM performance in chimpanzees (e.g., Carruthers, 2013). There are other studies of ape cognition that we were unable to decompose effectively into computational WM demands, and so these have not been included in our review of experiments. Examples (by title) are “Chimpanzees coordinate in a negotiation game” (Melis et al., 2009), “Five-year olds, but not chimpanzees, attempt to manage their reputations” (Engelmann et al., 2012), and “Direct and indirect reputation formation in nonhuman great apes (*Pan paniscus*, *Pan troglodytes*,

*Gorillagorilla*, *Pongo pygmaeus*) and human children (*Homo sapiens*) (Herrmann et al., 2013).

For all the studies listed in Table 1, we have critically reviewed the experiment to see if it had design problems and whether the data from the experiment support the interpretations made of subject behavior. Our long-term goal is not just to be critical of the work of others, but to use informed critique to help explain and interpret when, where, and how working memory evolved uniquely human characteristics. This requires considering working memory not only from the viewpoint of extant human and non-human primates, but also as an external physical resource registered in the stone artefacts that are part of the Paleolithic record of our hominin ancestors (Beaman, 2010; cf. Miller et al., 1960: 65). Because great apes are phylogenetically close to humans, an important consideration is how humans, unlike the great apes, came to develop the ability to foresee the appropriateness of the affordances of raw materials for making artefacts that bear scant resemblance to the forms of the raw material (e.g., the foresight of seeing a hand-axe in a block of rock). Great apes seem to lack that capacity for prospective memory and the requisite imagination needed to envisage “how things might be as well as how they actually are” (Russon, 2004) and for “blending reality among different potential spheres” (Manrique and Walker, 2017, pp. 41–53). Thus, we also discuss experimental work on stone artifact manufacture that relates to the evolution of human working memory (see Section 8).

### 2.1. Association between working memory and cognitive performance

That measures of WM performance should be associated strongly with other measures of cognitive performance is implied by Baddeley and others' models for WM as a primary nexus for the computational aspects of the brain. Alloway and Alloway (2010) briefly refer to several studies showing the wide-ranging occurrence of the expected association:

Individual differences in working memory capacity have important consequences for children's ability to acquire knowledge and new skills (see Cowan and Alloway, 2008, for a review). In typically developing children, scores on working memory tasks predict reading achievement independent of measures of phonological skills (Swanson and Beebe-Frankenberger, 2004). Working memory is also linked to math outcomes; low working memory scores are closely related to poor performance on arithmetic word problems (Swanson and Sachse-Lee, 2001) and poor computational skills (Bull and Scerif, 2001; Geary et al., 1999). Working memory capacity also has a significant impact on learning in various developmental disorders such as reading disabilities (Gathercole et al., 2006), language impairments (Alloway and Archibald, 2008), and motor difficulties (Alloway, 2007).

Others report that there is an association of WM with fluid intelligence and academic performance (e.g., Cowan et al., 2005; Cowan et al., 2006; Engle, 2002; Engle et al., 1999). Recent research suggests that the association of WM with fluid intelligence may be the consequence of both WM and fluid intelligence requiring attention control, which involves “the domain-general ability to regulate information processing in service of goal-directed behavior” (Burgoyne and Engle, 2020, p. 624). For academic performance, WM is also a strong predictor (Gathercole et al., 2004; Dumontheil and Klingberg, 2012) and WM skills at 5 years of age are even a better predictor of academic performance 6 years later than is IQ (Alloway and Alloway, 2010). Further supporting the connection between WM and mental performance is the significant genetic covariance among IQ, processing speed and WM as shown through a twin study based on monozygotic and dizygotic twin pairs (Luciano et al., 2001).

In the reverse direction, low size of working memory correlates with poor academic performance in reading, mathematics and science

<sup>1</sup> Miller's  $7 \pm 2$  is not a statistical formulation such as *mean*  $\pm$  *sd* but instead expresses the range of values likely to be observed in practice for the holding capacity of the human working memory system. Cowan's  $4 \pm 2$  is an analogous expression but differs from Miller's  $7 \pm 2$  in that Cowan's expression stipulates the range of values typically observed in different individuals for the parametric capacity of the short-term memory component of the working memory system. In analogy with Miller's  $7 \pm 2$ , the senior author introduced (Read, 2008) the expression  $2 \pm 1$  for the working memory holding capacity of *Pan* inferred from qualitative differences in their nut cracking abilities.

**Table 1**  
Chimpanzee Experiments Relating to WM.

Experiment	Task	Publication	WM subsystem	Longitudinal	WM Claim	Comment
Battery of Cognitive Tests	Age for first passing cognitive test	Herrmann et al. (2007), Wobber et al. (2014) (see p. 18 ff.)	SWM, ATM	Yes (synchronic)	Chimpanzees perform less well in the SWM and ATM subsystems than humans and have shorter cognitive development time	Expand linear change in HWM with age to SWM and ATM
Nut Cracking	Tool making and tool usage	Matsuzawa (1994) (see p. 20 ff.)	HWM, SWM	Yes (diachronic & synchronic)	none	Data show HWM = 2 ± 1
Floating Peanut	Innovative solution	Hanus et al. (2011) (see p. 25 ff.)	SWM, ATM	No	Chimpanzees match human cognitive ability of human 7-year-olds	Chimpanzees are comparable to human 3.5-year-olds
Memorize Order of Digits	Digit recall	Kawai and Matsuzawa (2000) (see p. 28 ff.); Inoue and Matsuzawa (2007) (see p. 28 ff.)	HWM	No	Chimpanzee HWM > human HWM	Data show HWM = 2 ± 1 for chimpanzees
Hidden Food	Search for food in closed boxes without repeating any box	Völter et al. (2019) (see p. 30 ff.)	HWM, SWM	No	Chimpanzees have remarkable WM ability	Data show HWM = 2 ± 1 for chimpanzees
Rotating Paddles	Planning ability	Tecwyn et al. (2013) (see p. 46 ff.)	SWM	No	Shows limited planning abilities	Data belie claims about sophisticated planning abilities in chimpanzees
Anticipated Need	Select or make tool in anticipation of future need	Mulcahy and Call (2006) (see p. 47 ff.), Bräuer and Call (2015) (see p.48 ff.)	SWM	No	Act in accordance with future needs	Data show limited ability to make choices in anticipation of future need
Goggles	Follow human gaze as a function of previous experience with goggles	Karg et al. (2015) (see p. 48 ff.)	ATM	No	Ape has level-1 visual perspective taking abilities	Apes failed to follow gaze as a function of their own previous visual experience with goggle-like masks
Transparent versus Opaque Screen	Attribute mental states that one has experienced to others	Kano et al. (2019) (see p. 51 ff.)	ATM	No	Apes take into account the visual perspective of competitor (Theory of Mind)	Modest performance (circa 60% success)
Subitizing (Rhesus Monkeys)	Distinguish larger from smaller set of items	Hauser et al. (2000), discussed by Carruthers (2013) (see p. 40 ff.)	HWM	No	Size of subitizing = Size of pure HWM	Size of subitizing is not a measure of pure HWM
Analog Estimator (Chimpanzees)	Distinguish larger from smaller set of items	Beran and Beran (2004) (see p. 41 ff.)	HWM	No	n.a.	Hauser experiment does not measure pure HWM

(Gathercole, 2008) and can lead to difficulty in participating in conversations (Bayles, 2003). In addition, working memory impaired through a stroke is found to have a negative effect on daily life activities (Fitri et al., 2020).

### 3. Age-based working memory development

Comparison of the size of HWM for our species with that of the non-human primates involves more than just a comparison of adult HWM capacity since, in humans, working memory becomes active at least as early as 7 months (Diamond and Doar, 1989) and then increases in size to  $7 \pm 2$  – the characterization for the size of adult human WM popularized by George A. Miller (1956) – in puberty or shortly thereafter. Thus, our first task is to delineate how HWM changes in size in humans with age. Then we can compare the development of HWM in humans with that of chimpanzees.

We begin the first task with a diachronically measured longitudinal study of human infants showing that HWM increases linearly with age. We use the result of that study to predict a linear pattern for the increase in HWM through puberty. Next, we validate and then corroborate the predicted linear pattern. We did the validation in two ways. First, we showed that the age-based trajectory for children from 6 to 12 years of age with HWM capacity measured both diachronically and synchronically matches both the linear pattern and the growth rate of the predicted trajectory. Second, we did a meta-study based on 12 synchronically measured data sets involving a variety of measures for working memory, with at least 3 age cohorts in each study so that there are at least 3 data points for inferring the trajectory pattern for the increase in HWM with age. We found that the trajectory for each of the 12 studies is linear and has the same growth rate as does the predicted trajectory. Lastly, we corroborate the validated predicted trajectory by

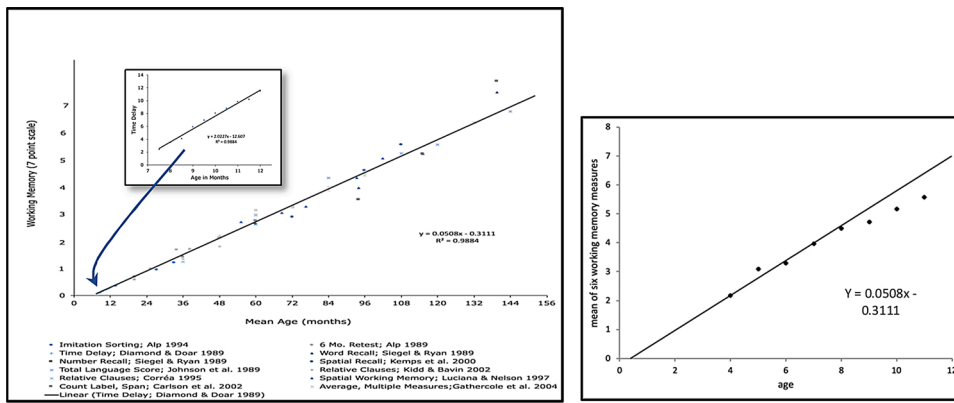
applying it to a study showing the longitudinal pattern for HWM increase in each of the components of the Baddeley model for working memory inferred from synchronic data on age cohorts varying from 4 to 12 years of age. For each component and measure of working memory, we find that the trajectory is linear with a growth rate matching that of the predicted trajectory. With this validation and corroboration of the predicted trajectory, we have a baseline pattern for increase in HWM by age in humans.

#### 3.1. Determination of the trajectory for age-based increase in HWM capacity

Studies relating the size of HWM to age longitudinally generally measure HWM synchronically across different age cohorts and then compare the HWM measurements for the different age cohorts to infer a longitudinal trajectory. One study by Diamond and Doar (1989), however, directly measured longitudinal changes in HWM by age in infants, starting with a cohort of 7-month-old infants and then measuring HWM monthly until the infants reached an age of 12 months. They used the length of a delayed response as their measure of working memory capacity. Their study shows a strong linear increase in HWM capacity for the infants:  $\text{time delay} = 0.16 * \text{age}_{\text{months}} - 1.01$  ( $r^2 = 0.99$ ) based on the mean working memory score for each age group (see Fig. 1A). Read (2008) extended this linear relationship through puberty to form a predicted linear trajectory for the increase in HWM capacity with age (see Fig. 1B).

#### 3.2. Validation of the extension of the linear infant growth trajectory beyond infancy

Next we validate the extension of the linear pattern for the age-



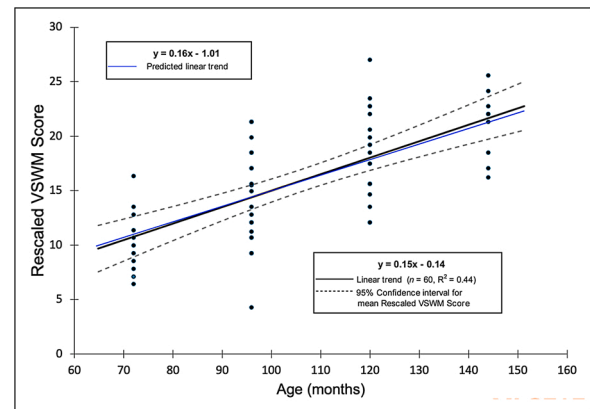
**Fig. 1.** A (Inset, left diagram): Growth pattern based on a diachronic longitudinal study from 7 to 12 months in age of the size of working memory for infants (Diamond and Doar, 1989). B (Left diagram): Meta-analysis of 12 synchronic longitudinal studies measuring change in working memory with age. Each data set is rescaled so that (mean age, mean WM size) for the data set is on the projected growth pattern for working memory from the Diamond and Doar study (see inset). The regression line for each data set has statistically the same slope as the projected growth pattern (see Read, 2008). The regression equation is the inset regression equation rescaled for age 144 months to correspond to size of HWM = 7. C (Right diagram): Size of working memory versus age compared to the projection of the Diamond and Doar (1989) growth trajectory, with working mem-

ory measured by the average in each age group of the 6 measures of WM discussed in Gathercole et al. (2004). (Alp, 1989; Alp, 1994; Carlson et al., 2002; Corrêa, 1995; Johnson et al., 1989; Kidd and Bavin, 2002; Siegel and Ryan, 1989).

related increase in infant HWM capacity beyond infancy. Validation of the linear extension of the infant trajectory makes use of a study conducted by Henrik Ullman and colleagues (Ullman et al., 2014) to determine the extent to which magnetic resonance imaging (MRI) and HWM capacity measurements obtained in one year (hereafter year 1) can predict the measurement of working memory capacity two years later (hereafter year 2). The HWM measurement used by Ullman and colleagues was a composite score based on three measures of Visuospatial WM (VSWM). Their study used a sample of  $n = 60$  children divided into age cohorts spaced two years apart, starting with a cohort of 6-year-old children and ending with a cohort of 20-year-old young adults. Thus, their dataset includes both diachronic data in the form of HWM measurements made 2 years apart and synchronic longitudinal data in the form of age-based cohorts of children, also 2 years apart.

In their analysis of the measurement data, Ullman and colleagues fit a curvilinear model to the relationship between age and their composite VSWM measurement, with the age of the children on whom the measurements were made varying from 6 to 20 years. They did not include statistical validation for the fit of the curvilinear model to the data. Consequently, their fitting of a curvilinear model to their data is not indicative of the data pattern from 6 to 12 year of age.<sup>2</sup> The slowing growth rate beyond the 12-year-olds makes the overall data pattern from 6 to 20 years appear curvilinear regardless of the trajectory pattern for ages 6–12. For this reason, we excluded children with age > 12 years in the Ullman data set so as to not include children with HWM capacity growth decreasing due to maturation. In addition, we excluded children who were 12 years old in year 1 to ensure that all the children in the reduced data set were each measured twice: once in year 1 and then again in year 2.

The relationship between age and the VSWM scores for the reduced data set is shown in Fig. 2. The residuals for the linear trendline (black line in Fig. 2) are normally distributed (Shapiro-Wilk Score = 0.99,  $p = 0.79$ ), hence a linear trend line cannot be rejected. The predicted trendline based on infant data is shown in blue and visually it is evident



**Fig. 2.** Graph of Visuospatial Working Memory (VSWM) versus age. Linear fit to data shown by the black line with a 95 % confidence interval for the data means. The blue line shows the predicted linear fit based on the Diamond and Doar (1989) infant data set (see Fig. 1A). The VSWM variable has been rescaled so that the point (mean VSWM, mean age) is on the predicted linear fit. The linear fit to the VSWM data (black line) has a slope of 0.15 in comparison to the slope of 0.16 for the predicted linear fit (blue line). Visually the predicted linear fit falls well within the 95 % confidence interval, hence the two lines are statistically identical. VSWM data were determined from Fig. 1 in Ullman et al. (2014).

that the slopes of the two trendlines are virtually identical, thus the growth rate in HWM for infants continues to be the growth rate in children from 6 to 12 years of age, even though the infant data are based on Auditory WM whereas the Ullman data are based on Visuospatial WM. We find, then, that these and the Ullman and colleagues data set validates the predicted linear trend line for children beyond infancy.

### 3.3. Corroboration of baseline predicted trajectory with a meta-analysis of 12 synchronic longitudinal studies

The predicted trajectory was also compared to the trajectory determined from the meta-analysis of 12 data sets based on age cohorts (Read, 2008). These two trajectories are fully consistent with each other (see Fig. 1A, B), thus unequivocally supporting the hypothesis that working memory increases linearly with age in *Homo sapiens*.

For the meta-analysis Read culled 12 studies – conducted between 1989 and 2004 and mainly from the psychological literature – that measured the size of HWM capacity using synchronic longitudinal data sets. A variety of keywords were used to retrieve the data sets, including

<sup>2</sup> Visually, it appears that the Ullman and colleagues’ data set shows a linear change in their measure of WM with age (see Fig. 4 in Ullman et al., 2014). For this reason, the senior author asked Ullman whether his data were linear over the age range from 6 to 14 years. He replied: “I agree with your interpretation of the data set we used in the study; WM development within the 6–14 years old interval appears linear. This is in concordance with other similar data sets. ... non-linear functions would not show a significantly better fit compared to a linear function if performed on this subgroup. ... For practical purposes we can therefore consider WM development in the age interval as linear” (Ullman, 2019).

both more obvious keywords (e.g., working memory, short term memory, age range, age cohorts, executive function, etc.) and keywords focusing on what was being measured (e.g., relative clauses, imitation sorting task, word memory, counting, etc.). To be included, a study had to measure HWM capacity over at least three age cohorts to provide enough data points to be able to determine whether the change in HWM capacity across the age cohorts in that study varies linearly with age. All studies that met these criteria were included in the meta-analysis. Jointly, the 12 studies cover the age range from 4 years through puberty. Each study shows a linear increase in the size of HWM capacity with age.

The 12 studies in the meta-analysis were each fit to the predicted linear growth pattern by rescaling the working measure capacity used in a study so that the point defined by (weighted mean age, weighted mean HWM capacity) is on the predicted growth trajectory. For each study, the growth rate of the HWM capacity measure matches the [Diamond and Doar \(1989\)](#) predicted growth rate for HWM capacity ([Read, 2008](#); see [Fig. 1B](#)). The meta-analysis also shows that all 12 of the data sets have the same growth rate for HWM capacity even though different measures of HWM capacity were used and the studies differed regarding the components of working memory hypothesized by [Baddeley](#) that were being measured. Altogether, these data validate a linear pattern for the growth of HWM capacity with age.

#### 3.4. Corroboration of the predicted baseline trajectory based on a large synchronic longitudinal study

Our corroboration of the predicted linear trajectory is based on a large study ( $n = 366$ ) conducted by [Gathercole and colleagues \(Gathercole et al., 2004\)](#) showing how a measure of working memory capacity for each of the executive, space and phonological components of [Baddeley's](#) model varies with age. They inferred the pattern for longitudinal change in each of the three WM components with a synchronic study using 9 age cohorts ranging from 4 to 12 years of age and with cohort sizes varying from 37 to 101. They found a strong linear increase in working memory capacity with age for each of these three components (see [Fig. 3](#)):

The developmental functions for measures associated with the phonological loop, the central executive, and the Visuospace sketchpad were found to be very similar, showing linear increases in performance from 4 years through to adolescence [and] ... for verbal storage only,  $y = 242x - 1.275$ ,  $r^2 = 0.971$ ; complex memory span,  $y = 269x - 1.521$ ,  $r^2 = 0.969$ ; Visuospace memory,  $y = 251x - 1.453$ ,  $r^2 = 0.979$  [pp. 187, 180–181].

The [Gathercole et al. \(2004\)](#) data show that the three measures for each WM component have strikingly similar linear patterns (see [Fig. 3](#)), with the exception that (1) the phonological loop shifts to a lower growth rate after 6 years of age and (2) the rate of growth in HWM capacity slows down and stabilizes at 11 years of age for the central executive and Visuospace components. Taking these observations into account, we compared the 6 measures comprising the central executive and Visuospace components for ages from 4 to 11 to the [Diamond and Doar \(1989\)](#) predicted growth trajectory by rescaling each of the [Gathercole et al. \(2004\)](#) measurements in the manner discussed above and then the six rescaled measurements were averaged within each age group to form a single age varying summary measure. The predicted growth trajectory (see [Fig. 1B](#)) fits the growth pattern for the summary measure for the data points from 4 to 8 years of age followed by a slight decrease in the growth rate of working memory capacity starting at 9 years of age (see [Fig. 1C](#)).

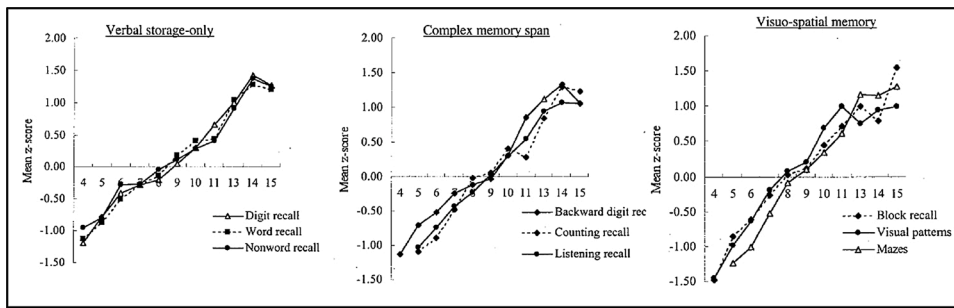
#### 4. Parallel between working memory development and prefrontal cortex development

Though working memory, in its totality, may be associated with more than just the prefrontal cortex, the executive function of working memory is regarded frequently as being a cognitive aspect attributable to prefrontal cortex (cf., [Fuster, 2002](#); [Goldberg, 2002](#); [Funahashi, 2017 and references therein](#)). The precise linkage is subject to different interpretations, with one interpretation being that the prefrontal cortex is where the information processed by working memory is stored (see [Lara and Wallis, 2015 and references therein](#)) and another that the role of the prefrontal cortex is to “select information and perform executive functions that are necessary to control the cognitive processing of the information ([Postle, 2006](#))” ([Lara and Wallis, 2015, p. 5](#)). In either case, the linkage of WM with the prefrontal cortex suggests that another way the pre-adult development of working memory performance with age can be measured is by taking advantage of both the fact that the pre-frontal cortex does not fully develop until adulthood ([Gogtay et al., 2004](#)) and the development of WM performance with age parallels the maturation of the prefrontal cortex ([Spear, 2000](#)) and the frontal lobes ([Conklin et al., 2007](#)) with age. Regarding the cortex, “the sequence in which the cortex matures parallels cognitive milestones in human development” ([Casey et al., 2005:104 and references therein](#); see also [Fuster, 2002](#)). In addition, maturation of the connectivity between the prefrontal region and the parietal region also correlates positively with increases in WM performance ([Olesen et al., 2003](#)).

Developmental increase in the prefrontal cortex has been measured both by changes in grey and white matter volumes and myelination of axon fibers ([Casey et al., 2005](#); [Zhou et al., 2016, and references therein](#)) and by increase in prefrontal cortex activity with the cranial location of the activity measured using fMRI ([Casey et al., 2005](#)). Regarding the association between WM maturation and maturation of the prefrontal cortex, there is a positive association between increase in WM performance and maturation of white and gray matter ([Darki and Klingberg, 2015](#)).

[Klingberg et al. \(2002\)](#) observe that previous studies either documented increases in WM capacity with age until adulthood or increases in neuronal activity with age but not correlations between WM capacity increase and neuroactivity increase. Their study was designed to fill this gap. For  $n = 14$  children with ages ranging from 9 to 18 years, they measured WM capacity with a visuospatial task consisting in locating the previous position of three to nine circles sequentially flashed on screen, and measured brain activity using fMRI scans while the WM task was being performed. This allowed them to compute “correlations between brain activity and age, as well as between brain activity and WM capacity” (p. 1), with these correlations reported in their article as either positive, negative, or not significant. Positive correlations between age and WM capacity, and between age and activity, were found in the superior frontal and intraparietal cortex and interpreted as showing a “positive correlation between age related increases in WM capacity and brain activity in the superior frontal and intraparietal cortex” (p. 8).

More generally, the study reported in [Conklin et al. \(2007\)](#) establishes that the temporal order in which performance was mastered for the WM tasks that were part of their study parallels the temporal order for the developmental maturation of the frontal substrates –i.e., the ventrolateral prefrontal cortex and the dorsolateral prefrontal cortex – associated with the performance of these WM tasks. Altogether, it appears that “cortical activity and structure are closely related to current WM, while the subcortical white matter tracts and activity in the caudate are preceding these changes and predict future WM capacity” ([Darki and Klingberg, 2015: 1594](#)), hence WM capacity closely relates to brain development and maturation associated with age.



**Fig. 3.** Three measures of Holding WM capacity, grouped by task type, for each of Baddeley’s WM components (Verbal storage, Complex memory span, Visuo-spatial memory) using z scores versus age for each measure. Copyright © 2004 by American Psychological Association. Reproduced with permission. Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40(2), 177–190. <https://doi.org/10.1037/0012-1649.40.2.177>.

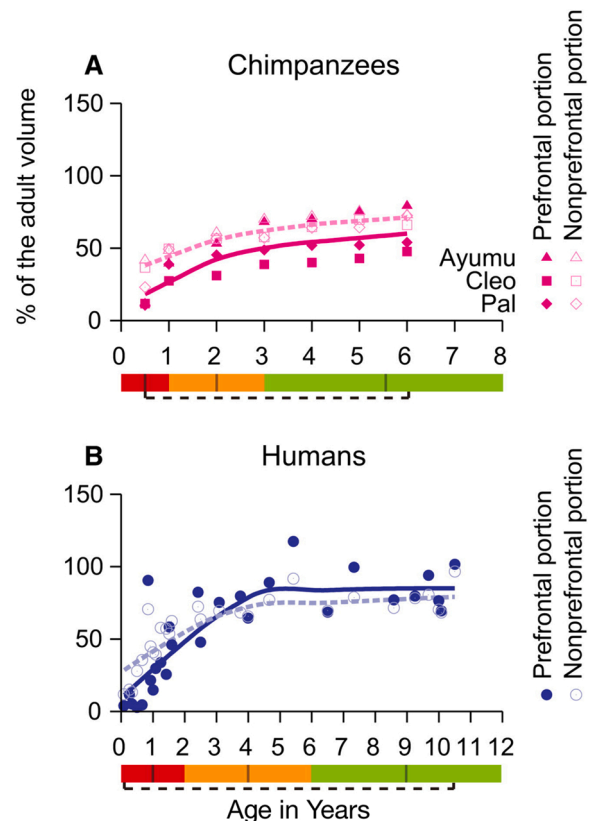
**5. Cognitive development in humans parallels cognitive development in non-human primates**

In general, the non-human primates have cognitive abilities and brain functions that are like those of humans (King et al., 1988), including working memory. Within working memory there is a close parallel between the Visuospace component of working memory for humans and for non-human primates such as monkeys and the great apes. Though working memory of humans and non-human primates show many parallels, there are differences, such as the ability of humans to actively maintain information in working memory through means such as rehearsal using the “phonological loop” (Baddeley, 2003). As discussed by Brady and Hampton (2018), it has been assumed that while rhesus monkeys can keep familiar stimuli in working memory through rehearsal, the same is not the case for novel stimuli since they lack the language facility to enable the verbal recoding and rehearsal through a phonological loop used by humans to keep novel stimuli active in working memory. Their recent research, however, shows that rhesus monkeys can keep novel stimuli active in working memory through encoding that does not depend on language, thus opening the possibility that “mechanisms supporting the maintenance of unfamiliar memoranda in working memory evolved before language and at least 32 mya when a common ancestor of humans and rhesus monkeys lived” (p. 8).

Working memory, then, facilitates comparison of the cognitive development in the non-human primates with that of humans given the association, in both groups, with the maturational pattern of the prefrontal cortex. The pattern of increase in WM performance with age in humans also occurs in monkeys, with WM performance already in place before 1 year of age (Alexander and Goldman, 1978; Diamond, 1990) and increasing with age to 36 months and then through puberty until early adulthood (Zhou et al., 2016). For the rhesus monkeys (*Macaca mulatta*), an experiment carried out by Goldman and Alexander (Goldman and Alexander, 1977; Alexander and Goldman, 1978) shows that there is a parallel between increase in WM performance with age and increase in prefrontal cortex development and activity from 12 to 36 months. They measured, at different ages, their performance on a 12 s delayed response task, both under normal conditions and when the activity of the dorsolateral prefrontal cortex (DLPFC) had been depressed by induced hypothermia. From the difference between these two measures, they could determine the magnitude of the effect of the DLPFC reduced activity on the performance level for the delayed response task. They found no effect in the 9 – 16-month-old monkeys, a reduction of 7–8% in the 19 – 31-month-old monkeys, and a reduction of 21–25 % in the 36-month-old monkeys. This development parallels the pattern found in humans from adolescence to adulthood: “Our findings establish non-human primates as a model of cognitive development that mirrors the progression of working memory ability observed in humans during adolescence” (Zhou et al., 2016: 7).

**5.1. Cognitive development in humans compared to cognitive development in non-human primates**

When compared to human infants and using the same task for both human infants and monkey infants, macaque performance at 4 months matches human toddler performance at 12 months (Diamond, 1990), indicative of the development extension that is characteristic of humans (Montagu, 1955). The protracted maturation of the prefrontal cortex in humans relates, it has been suggested, to neuronal plasticity that allows for increased effect of post-natal experiences on brain functioning (Johnson, 2001). It has also been suggested that the same extension in time to maturation occurs with the chimpanzees and so it would be present in an ancestor common to chimpanzees and humans (Sakai et al., 2012), though there are differences in prefrontal cortex maturation rates between humans and chimpanzees. Smaers et al. (2017: 720)



**Fig. 4.** Graph of the proportion of prefrontal white matter by age (in years). Red: infant, orange: juvenile, green: adult. Copyright © 2011 by Elsevier. Reproduced with permission. Sakai, T., Mikami, A., Tomonaga, M., Matsui, M., Suzuki, J., Hamada, Y., Tanaka, M., Miyabe-Nishiwaki, T., Makishima, H., Nakatsukasa, M., Matsuzawa, T., 2011. Differential prefrontal white matter development in chimpanzees and humans. *Current Biology* 21(16), 1397-1402. <https://doi.org/10.1016/j.cub.2011.12.052>.

demonstrated that “expansion of human prefrontal cortex significantly exceeds the enlargement in other heteromodal association areas, suggesting that human evolution has been characterized by selection for changes in executive functions mediated by this cortical region”.

### 5.2. Prefrontal cortex development rate faster in humans in comparison to chimpanzees

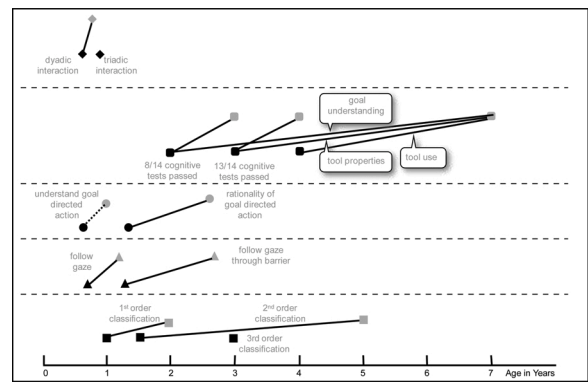
Prefrontal cortex develops at a faster rate in humans than in chimpanzees (Somel et al., 2011). As shown by Sakai and colleagues (Sakai et al., 2012), a longitudinal study of prefrontal development in chimpanzees (*Pan troglodytes*) from 6 months to 6 years using magnetic resonance image scans shows that the prefrontal white matter volume is still immature and has not reached adult levels during prepuberty (see Fig. 4A). In sharp contrast, in humans (*Homo sapiens*) prefrontal development extends from 1 month to 10.5 years and is part of the development extension that is characteristic of our species (Montagu, 1955). Adult prefrontal levels are reached well before puberty (see Fig. 4B). This protracted maturation of the prefrontal cortex in humans relates, it has been suggested, to neuronal plasticity that allows for increased effect of post-natal experiences on brain functioning (Johnson, 2001). The rate of prefrontal white matter volume increase is slower during infancy in chimpanzees than in humans, thus suggesting “a less mature and more protracted elaboration of neuronal connections in the prefrontal portion of the developing brain existed in the last common ancestor of chimpanzees and humans” (Sakai et al., 2012).

### 5.3. Differences in development of cognitive capacities in humans in comparison to chimpanzees

#### 5.3.1. Developmental delay in *Pan* versus *Homo* accelerates with *Pan* maturation

There have been two major studies that compare the cognitive abilities of humans and chimpanzees. Herrmann and colleagues (Herrmann et al., 2007) compared humans, chimpanzees, and orangutans and Wobber and colleagues (Wobber et al., 2014) compared humans, chimpanzees, and bonobos by applying a battery of cognitive tests that were divided into social cognition versus physical cognition tests. The first study, by Herrmann and colleagues, compared 105 toddlers (mean age of 2.5 years, varying from 2.3 to 2.7 years) from a medium size city in Germany to 106 chimpanzees from the Ngamba Island (Uganda) and Tchimpounga Chimpanzee Rehabilitation Centre (Republic of Congo) sanctuaries, with chimpanzee ages varying from 3 to 12 years (mean age = 10 years, standard deviation = 5.20 years<sup>3</sup>). The study compared the two groups as a whole, not with a time trajectory for the individual chimpanzees. Thus, toddler humans were compared to mainly adult chimpanzees. They found that the human toddlers significantly ( $p < 0.001$ ) outperformed the chimpanzees on measures of social cognition but did not outperform the chimpanzees for measures of physical cognition (Herrmann et al., 2007: Fig. 1).

The second study is by Wobber and colleagues and expands on the scope of the first study by comparing the cognitive development pattern in humans to that of chimpanzee and bonobo “toddlers,” with the individuals in each group given 14 cognitive tasks relating to social skills, physical skills, and attentional/motivational control. The chimpanzees performed better than the bonobos at understanding physical causality and the bonobos performed better than the chimpanzees at understanding social causality, thus showing “the role of ecological and socio-ecological pressures in shaping cognitive skills over relatively short periods of evolutionary time” (Herrmann et al., 2010, p. 1), rather than a consistent difference in WM between chimpanzees and bonobos. For this reason, the bonobos will be subsumed with the chimpanzees for our



**Fig. 5.** Age for task emergence for human and chimpanzee toddlers. Age for task emergence is defined as the age at which a task is first performed successfully by > 50 % of group members. Black symbols — human toddlers; grey symbols — *Pan* “toddlers.” Lines connect similar task for human toddlers and for *Pan* “toddlers.” Tasks from top to bottom, with kind of task in parentheses and followed by references, are:

Dyadic to Triadic Interactions (social cognition). Note: True triadic interaction (infant, mother, object) has not been shown to occur in *Pan*. Carpenter et al. (1998), Carpenter and Call (2013), Tomasello (1999), Hayashi and Matsuzawa (2003) and Nakashima (2003).

Panel of Social and Physical and Cognitive Tests. Tool properties (physical cognition), tool use (physical cognition) and goal understanding (individual to social cognition) are for single tests from the panel of cognitive tests and are the most difficult tests for *Pan* “toddlers.” Task emergence: (1) 67 % of social and 60 % of physical cognition tests emerged in 2-year-old humans; (2) 44 % of social and 0% of physical cognition tests emerged in 2-year-old chimpanzees; (3) 100 % of social and 80 % of physical cognition tests emerged in 3-year-old humans; (4) 55 % of social and 20 % of physical cognition tests emerged in 3-year-old chimpanzees. Wobber et al. (2014).

Goal Directed Action (social cognition). Gergely et al. (2002), Woodward (1998), Bering et al. (2000), Tomasello and Carpenter (2005) and Wobber et al. (2014).

Gaze Following (social cognition). Moll and Tomasello (2004), Okamoto-Barth et al. (2008); Tomasello and Carpenter (2005) and Wobber et al. (2014).

Spontaneous Classification. Note: 3rd order or higher classification does not occur in *Pan*. Langer (1986), Poti et al. (1999) and Spinozzi et al. (1999).

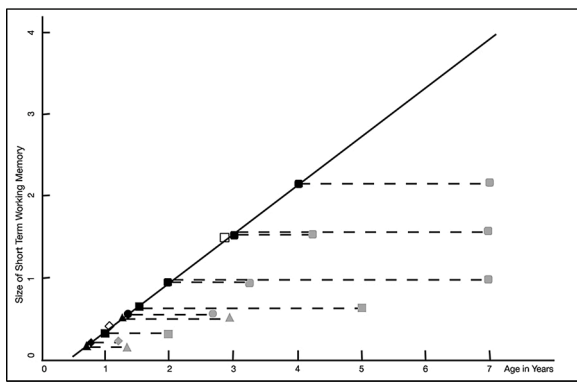
purposes here. The human sample consisted of  $n = 48$  children from age 2–4 and the comparison chimpanzee sample consists of  $n = 49$  chimpanzees also from age 2–4. A second sample of  $n = 44$  juvenile chimpanzees were tracked longitudinally for 3 years.

The authors begin by observing that for the first few months after birth, cognitive capacities in human and ape infants are similar (see also Langer, 2006; Matsuzawa et al., 2006; Rosati et al., 2014). For the cognitive tasks, the researchers found that from age 2–4 years, human toddlers outperform *Pan* “toddlers” with respect to social cognitive skills at all ages (see Fig. 5) and also develop increased social cognitive skills with age whereas *Pan* “toddlers” do not. In the physical domain, the performance of human toddlers and ape “toddlers” are comparable at two years of age but then increasingly differed as human toddlers augment their level of performance with age, but *Pan* “toddlers” do not. For the attentional/motivational control tests, there is no difference in performance between human toddlers and *Pan* “toddlers” at any age and neither increases in performance with age (Wobber et al., 2014). From 4 to 6 years of age, cognitive performance of *Pan* subjects increases, but at a slower rate than is the case for the human subjects from 2 to 4 years of age (Wobber et al., 2014; see Fig. 5). The researchers conclude:

there are significant differences in the pattern and pace of cognitive development between humans and ... chimpanzees and bonobos. First, ... accelerated ontogeny in human cognitive development relative to other apes.... Second, ... variable pattern of cognitive development between human and *Pan* ... particularly apparent within the social cognitive domain.... [H]uman cognitive

<sup>3</sup> The authors do not include the standard deviation. We computed it from Table S1 in the Supplemental Data for their article.





**Fig. 6.** Age when task or test can be first performed. Data are from Fig. 5. (See Fig. 5 for data references.) Black symbols — human infants; grey symbols — Pan infants; open symbols — occurs in human infants but not Pan infants. Dashed lines connect similar task or test for human infants and Pan infants. Solid line: Predicted change in size of WM with age in humans (from Fig. 1B). Assuming the same size of WM is required for doing the same task or test in Pan as in *Homo sapiens*, the data show a delay in the age for first performance of the same task or test in Pan infants in comparison to human infants. This delay implies a slower growth rate for WM in Pan versus *Homo sapiens*. Chimpanzees reach puberty by age 7. These data imply that the size of WM is, on average,  $\sim 2$  for Pan adults.

◆ — Dyadic to Triadic Interactions; ■ — Panel of Social and Physical and Cognitive Tests;  
● — Goal Directed Action; ▲ — Gaze Following; ■ — Spontaneous Classification

development is accelerated overall in comparison to our closest living relatives, with particularly marked change in humans' development of socio-cognitive skills. (pp. 567–568; see Fig. 5)

### 5.3.2. Delays in occurrence of comparable cognitive abilities in Pan implies a limited working memory size for Pan in comparison to humans

The pattern derived from the Wobber et al. (2014) data is that cognitive capacities of human toddlers and Pan “toddlers” are comparable, but the development of cognitive capacities in human toddlers occurs faster than in Pan “toddlers,” not only for the cognitive tests they used in their research (see middle part of Fig. 6), but for a variety of other behaviors relating to cognitive capacity as well (see upper and lower part of Fig. 6). For all data, an initial one-year delay for Pan infants to display simpler behavior comparable to human infants (leftmost symbols linked in pairs in Fig. 5) is extended to a 3–5-year delay for more complex behavior (rightmost symbols linked in pairs in Fig. 5) as shown in Fig. 6.

In the bottom row of Fig. 5 are summary results from the research work performed by Langer (1986) and Luciana and Nelson (1998) on the way undirected human infants and toddlers increasingly formulate more complex groupings of objects with age, going from 1st to 3rd order classifications. The research protocol used by Langer has also been applied to Pan by Potì et al. (1999) and Spinozzi et al. (1999). For Pan, 2nd order classifications are first made four years after making 1st order classifications and Pan “toddlers” never make 3rd order classifications. The failure of Pan to ever exhibit a behavior engaged in by human toddlers also occurs with interaction behavior (shown at the top of Fig. 5).

Dyadic interactions occur at a slightly older age in Pan “toddlers” than in human toddlers through “primary intersubjectivity” (Trevarthen, 1979), where infants attune their emotions to those of their caregivers through “face-to-face” exchanges (i.e., interlocking gazes). While human toddlers then begin, at 9 months (Tomasello, 1995), to engage in triadic interactions involving what Carpenter and Call (2013) refer to as joint joint attention in which the toddler is engaged with an object and with her/his caregiver while simultaneously the caregiver is

engaged with the toddler and the object, and each is aware of, and engaged in, the other's attention. However, “there is so far little if any convincing evidence that chimpanzees ... engage in truly joint joint attention” (Carpenter and Call, 2013: 51).

Overall, the pattern is clear. The first occurrence of a cognitively based behavior occurs at a later age in Pan “toddlers” and development beyond first occurrence is more rapid in human toddlers than in Pan “toddlers.” For some behaviors, such as 3rd order classifications or triadic interactions, the behaviors either do not occur in Pan, or only occur sporadically at best. Given the association between WM development and development of prefrontal cortex, and the association of cognitive performance with development of prefrontal cortex, it follows that WM development in Pan is limited in comparison to WM development in humans. The pattern for the occurrence of behaviors that relate to cognitive abilities parallels the pattern for the development of the prefrontal cortex in humans in comparison to Pan. Given the correlation between the development of the prefrontal cortex and the development of WM in humans, it follows that the pattern for the appearance and rate of growth in cognitive abilities in human toddlers is paralleled by equivalent growth in WM capacity. By the same token, the growth in WM capacity for Pan “toddlers” must be slower than it is for human toddlers. When we map the data shown in Fig. 5 onto the regression line for increase in the size of HWM with age in *Homo sapiens* (see Fig. 6), and we assume that the same cognitive performance requires the same size of HWM in Pan as it does in *Homo sapiens*, then the size of HWM for Pan by age 7 – that is, by the time of Pan reaching puberty – is  $2 \pm 1$  (using a characterization of HWM for Pan that parallels Miller's characterization for human HWM), far below Miller's  $7 \pm 2$  value for the size of HWM in adult humans.

## 6. Experimental data relating to holding working memory capacity in simiiform primates

### 6.1. Chimpanzee nut cracking performance

Perhaps the strongest natural experimental evidence for the small size of the effective working memory for chimpanzees is the inability of 25 % of the chimpanzees at Bossou to perform a three-part sequence required for cracking nuts, despite watching, day-in and day-out, other chimpanzees do the task successfully (see Matsuzawa, 1994 and other references in Read, 2008). The task requires selecting a stone anvil, putting a nut on the anvil, then hitting the nut on the anvil with a stone used like a hammer. We can characterize the sequence by (anvil, nut, hammer). It is reported that 25 % of the chimpanzees never learn to do this sequence. They either do (anvil, nut) and hit the nut on the anvil with a fist, meaning that they just conceptualize the (anvil, nut) part of the sequence, or put the nut on the ground and hit the nut with a stone, meaning that they conceptualize just the (nut, hammer) part of the sequence. In other words, though these chimpanzees understand and desire the nut meat, and even though they see other chimpanzees do the sequence (anvil, nut, hammer), they cannot replicate the sequence despite extensive trials and repeated observations of other chimpanzees doing this sequence.

Similarly, in a group of 16 chimpanzees (*Pan troglodytes*) released on an island off the coast of Liberia, 13 of the chimpanzees learned to crack nuts and 3 did not (Hannah and McGrew, 1987). In a different experiment (Marshall-Pescini and Whiten, 2008), 4 out of 12 chimpanzees failed to learn to crack nuts. For these two experiments, altogether 7 out of 28, or 25 %, failed to learn to crack nuts, matching the failure rate observed at Bossou. In yet another experiment, 13 chimpanzees in a zoo enclosure were provided with the materials needed to crack nuts (nuts, wood hammer and wood anvil) and went through a sequence of learning trials aimed at determining the order in which chimpanzees learned the steps required to crack nuts (Needle et al., 2020), but none of the chimpanzees learned to crack nuts. The authors attributed the failure to learn to crack nuts to “a certain level of developmental prowess [is]

required to express nut-cracking ... between maturation effects of the body and the brain” (p. 20), but the chimpanzees were already older adults and so beyond a critical, age-based learning “window” (Inoue-Nakamura and Matsuzawa, 1997).

Carruthers (2013) dismisses the data from Bossou by claiming, but without providing evidence, that it is due to chimpanzee “lack of understanding of physical forces and their effects” (p. 10373). However, at Tai National Park in Côte D’Ivoire, where exposed, tree roots are used as anvils instead of loose stones, all chimpanzees learn to crack nuts (Boesch and Boesch, 1983), hence, according to Carruthers’ claim, all these chimpanzees must have learned the “physical forces and their effects” but somehow this is not the case for the chimpanzees at Bossou or in the experiments with chimpanzees learning to crack nuts. The matter, though, is much simpler than whether all chimpanzees develop mental physics models in one region but not in another. Enacting the sequence (anvil, nut, hammer) does not require a mental physics model on the part of a chimpanzee, but just that HWM becomes populated by the objects manipulated by other chimpanzees to crack nuts and then imitating their actions. The simplest explanation for the 25 % that fail to crack nuts is that they have HWM = 2 and cannot hold simultaneously in HWM the full sequence (anvil, nut, hammer) (see discussion and references in Read, 2008), whereas at Tai National Park the exposed roots are permanent fixtures in the landscape, and so, following the argument of Woodman et al. (2007) that visually constant objects need not be initiated into WM in order to be active in visual memory, the tree roots need not be initiated into WM to carry out the nut cracking task. Hence for this group of chimpanzees, HWM = 2 suffices to crack nuts, thus all chimpanzees at Tai National Park can crack nuts.

### 6.2. Limited visuospace working memory in simiiform primates

Especially enlightening regarding the size of holding working memory (HWM) is a comparison of the Visuospace working memory in humans, chimpanzees, orangutans, and Rhesus monkeys required to match from 2 to 6 pairs of images hidden on several cards lying face-down (Washburn et al., 2007). Only the backs of the cards were displayed on a screen and the subject had to manipulate a joystick to move a cursor over a card so that its hidden image then appeared briefly. While it was still visible the image had to be consigned to working memory in order that it could be matched with its hidden, paired card image that might be uncovered subsequently. The hypothesis being tested is that the larger the number of different images, one from each pair of images, held in working memory while waiting for a matching one to appear, the greater should be the load imposed on working memory. The comparison procedure used clip art images, dot images, and lexigrams used in language training of chimpanzees. Performance varied according to the kind of image. Abstract dot images were found to

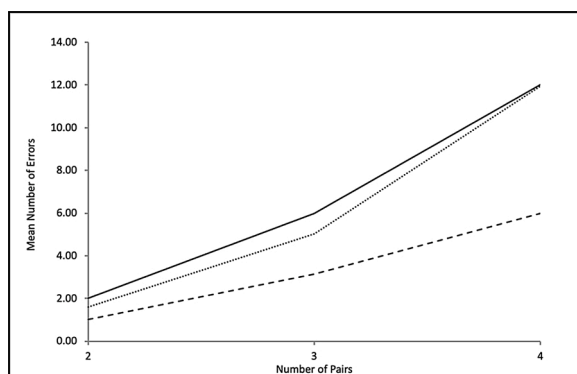


Fig. 7. Relationship between number of pairs of dot images and mean number of errors for *Pan* (dotted line) and *Homo sapiens* (dashed line). Mean number of errors for random choices shown by a solid line. Data for *Pan* and *Homo sapiens* (solid discs) from Washburn et al. (2007).

be least affected by prior experience with kinds of images. As shown in Fig. 7, while chimpanzees performed slightly better than at a chance with 2 or 3 pairs of images, their performance decreased to a chance level with 4 pairs of images. In contrast, humans perform increasingly better than chance with increase in the number of images. These data indicate that *Pan* has a much-reduced HWM size in comparison to humans.

### 6.3. Innovation and working memory comparison between pre-pubescent *Pan* and human children

The floating peanut test (Mendes et al., 2007; Hanus et al., 2011) requires innovation on the part of the subject for its solution, whether the subject is a human child or a pre-pubescent chimpanzee. The goal of the test is to retrieve a peanut from a transparent tube too narrow for manual extraction. The tube cannot be manipulated and the only way the peanut can be retrieved is by the participant realizing that if enough water is put into the tube, the peanut will float to the top and can be retrieved. The experiment was done under two conditions: wet – the peanut is initially floating on water partially filling the tube – and dry – the peanut is initially at the bottom of a dry tube.

The task, however, does not directly involve working memory. Instead, a successful solution by chimpanzees has been interpreted as showing a high level of mental performance on their part, even greater than humans (see Hunt, 2020). This interpretation indirectly implies a working memory size comparable to that of humans due to the correlation between IQ and the size of working memory. However, review of the data from this experiment does not support this interpretation.

The experiment was performed with  $n = 44$  chimpanzees from both the Leipzig Zoo and the Ngamba Island Chimpanzee Sanctuary (Hanus et al., 2011). None of the  $n = 23$  chimpanzees that did the experiment under the dry condition succeeded in retrieving the peanut. Of the  $n = 21$  chimpanzees that did the experiment under the wet condition,  $n = 5$  ( $p = 0.24$ ) succeeded in retrieving the peanut by spitting water into the tube. The successful chimpanzees each got a mouthful of water from a water dispenser and then spat the water into the tube, repeating this process until the peanut floated to the top.

When the experiment was done with children (4, 6 or 8 years old) with  $n = 72$  and divided equally among the three age groups and the two conditions for the experiment, water can be poured into the tube from a pitcher. The success rate of 4-year-olds was  $p = 0.17$  for the wet condition and  $p = 0.00$  for the dry condition, which is comparable to the performance of the pre-pubescent chimpanzees. Successful performance by children increased with age. The success rate of 6-year-old children was  $p = 0.50$  for the wet condition – twice the success rate of the chimpanzees for the wet condition – and  $p = 0.33$  for the dry condition.

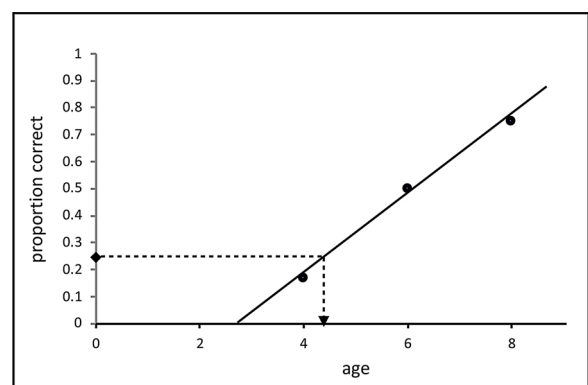


Fig. 8. Comparison of the performance of children (= ●), ages 4, 6 and 8, with that of chimpanzees (= ◆) in the floating peanut experiment. Linear fit to children's data is  $pc = 0.1458age_{years} - 0.4208$  ( $R^2 = 0.99$ ,  $p = 0.052$ ). Data are from Hanus et al. (2011).

The success rate for the 8-year-olds was  $p = 0.75$  for the wet condition – triple that of chimpanzees – and  $p = 0.42$  for the dry condition (Hanus et al., 2011).

These results indicate that the chimpanzees performed between the level of 4-year-old and 5-year-old children (see Fig. 8), hence they imply that the mean size of HWM for the *Pan* cohort is between 2.1 and 2.7 (see Fig. 1B), which is consistent with  $HWM = 2 \pm 1$  for chimpanzees and well below  $HWM = 7 \pm 2$  for humans.

#### 6.4. Advanced planning and size of holding working memory

In another experiment (Mulcahy et al., 2005), the subject had to choose between a short and a long stick-like tool to reach an otherwise out-of-reach grape reward. The conditions for the experiment varied from: the grape distance and the length of both tools are all simultaneously visible before a tool is selected by the subject, to: the grape and the two tools are each presented sequentially and each tool is hidden by a box after being presented so that the subject needs to keep in mind the distance of the grape and the length of each of the two tools in order to choose a tool that can retrieve the grape. (When the tools are hidden by a box, the subject selects a tool by pointing to a box and the experimenter removes the box to give the subject access to the tool.) In some of the experimental conditions both tools were suitable for retrieving the grape and in other conditions only the longer tool could be used to retrieve the grape. Under the simplest condition, with grape distance and both tools visible and both tools long enough to retrieve the grape, the subjects selected a tool randomly. When only the longer tool could be used to retrieve the grape, the longer tool was chosen preferentially. Under the more difficult sequential condition of the grape location being shown first and then both tools are shown simultaneously, the results were similar: the longer tool was selected preferentially when only the longer tool could be used to reach the grape. Otherwise, the choice of a tool was random. Finally, and critically, when the grape location and tools were each shown sequentially and each tool hidden by a box immediately after being shown, the longer tool was selected preferentially even when the shorter tool could be used to reach the grape.

Next, we want to determine what this experiment shows regarding the size of HWM. First, the preferential choice by a subject of the longer of the two tools presented to the subject is not accounted for by a strategy that makes minimal use of working memory resources. For example, a simple strategy like “When uncertain, only select the absolutely longest tool; otherwise, do not select a tool” does not account for their preferential behavior for the following reason. Assuming the subjects are familiar with the tools and know in advance which tool is the longest, and though this strategy makes little use of WM resources, it does not account for their preferential choice of the longer of two tools that are presented since the experiment involves three different tool lengths and the medium length tool is paired with the shortest tool in one half of the trials and with the longest tool in the other half of the trials. Consequently, by only selecting the absolutely longest tool, the subject would fail to select a tool half the time and would always select the longest tool half the time, whereas the subjects actually selected the medium length tool half the time and the longest tool the other half of the time. Their actual behavior meant that when the tools were presented sequentially (and each was hidden by a box after being presented), then to pick the longer of the two tools that were presented the subject “had to encode and remember the length of two tools separately (because no direct comparison was possible) and then mentally compare them in order to select the longest [sic.] one” (p. 31).

Next, to determine whether these results could be due simply to the subject not coding, for some reason, the distance to the grape when the grape and the tools are presented sequentially, another experimental condition involved sequentially first seeing the grape location and then just one tool. When only the shorter of two tools was presented, and if it was too short to reach the grape, the subject refused to select the shorter tool. Thus, the subject knew that it was not long enough to reach the

grape, meaning that the subject had encoded the grape distance in WM. Therefore, in the sequential trials three items had to be encoded in WM – the distance of the grape and the length of the two tools – if the behavior were to match the subject’s behavior when the grape and the two tools were presented simultaneously to the subject and so no information had to be encoded in WM. Hence, the subject’s divergent performance (preferring the longer tool regardless of the distance to the grape) in the sequential trials in comparison to being presented with the grape and the two tools simultaneously, indicates that the subject could not be keeping in mind all three of the distance to the grape and the length of each of the two tools. If the latter were the case, then the behavior with sequential trials would be same as the behavior with the grape and the two tools presented simultaneously. Instead, in the sequential trials the subjects appear to end up not remembering the distance to the grape and so, with uncertainty about the distance to the grape, preferentially selects the longer of the two tools even when the shorter tool can reach the grape. This indicates that the size of working-memory capacity is limited to two chunks of information. Hence, encoding the respective lengths of two tools in HWM requires deleting the grape distance from HWM and the subject has reverted to the strategy of simply selecting the longer tool when the distance to the grape is not known or is not remembered.

While this experiment involved orangutans and gorillas but not chimpanzees, the researchers found that there was no species difference in performance. Unless chimpanzees outperform the other great apes, the same results would be found for chimpanzees and the size of HWM would be 2 for *Pan*.

#### 6.5. Digit recall and size of working memory

An experiment conducted by Matsuzawa and colleagues (Matsuzawa et al., 2006), and with chimpanzee subjects already trained to recognize the ordinal sequence among single digit numbers, involved the display of 5 randomly selected single digit numbers on a computer screen. The subject chimpanzee touches the smallest of the numbers, and the other numbers are then masked by white rectangles. The chimpanzee now must touch the rectangles in the ordinal sequence of the hidden numbers. Since the first of the five numbers is touched while visible and since the last number is a forced choice, if there are  $n$  numbers then only the location of  $n - 2$  numbers must be recalled from memory to correctly identify the numbers in their ordinal sequence. The female chimpanzee Ai correctly identified the ordinal sequence 65 % of the time with  $n = 5$  numbers and  $k = 20$  trials (Kawai and Matsuzawa, 2000) which, at best, is equivalent to being correct 100 % of the time on the second and third numbers, then making a random guess for the third number 70 % of the time and remembering the location of the fourth number 30 % of the time. Statistically, her performance is consistent with the null hypothesis  $H_0: p = 0.5$  since  $B(k \geq 13 \text{ successes}, n = 20 \text{ trials}, p = 0.5) = 0.128$ , where  $B(k, n, p)$  is the binary distribution for  $k$  successes in  $n$  trials with probability  $p$  of a success, and so the null hypothesis is not rejected at the  $\alpha > 0.10$  significance level. Thus, her performance implies her size of  $HWM = 2$ . Kawai and Matsuzawa, however, erroneously assert she has size of  $HWM = 5$  since there are 5 masked numbers. They do not consider that the first number is not recalled from memory and the fifth one is a forced choice. Two other chimpanzees are correct 67 % and 62 % of the time, respectively with  $n = 4$  numbers (Inoue and Matsuzawa, 2007), hence perform less well than Ai and so their HWM size is at most 2. These data do not support the claim that the size of chimpanzee HWM is comparable to that of humans.

In another experiment carried out by Matsuzawa and a colleague (Inoue and Matsuzawa, 2007), up to 9 single digit numbers are randomly flashed on a computer screen for either 650, 430 or 210 ms, with the last below the time necessary for humans to switch from central to peripheral focus (Bartz, 1962). At this shortest interval, one chimpanzee, Ayumu, after extensive training, can point correctly to the ordinal order for eight single digit number locations on the screen with 80 % accuracy (Matsuzawa, 2009), but loses this capability when the

digits were flashed for 650 ms. This suggests that WM is not involved in Ayumu's performance, and the researchers attribute the performance of Ayumu to the long disputed (Gray and Gummerman, 1975) "eidetic imagery" (Jaensch, 1930) or to some form of camera-like memory (Matsuzawa, 2013: 444). Whether eidetic or camera-like, this is not working memory. Nonetheless, Matsuzawa incorrectly refers to the performance of Ayumu as exhibiting a HWM size exceeding that of humans and that the performance of Ayumu is not matched by humans (Matsuzawa, 2013). However, according to a meta-analysis of > 90 sets of data taken from delayed match-to-sample studies covering 25 species from bees to birds and mammals, including primates (Lind et al., 2015), recall performance without delay following a stimulus varies little among species due to a ceiling effect of almost perfect performance by each species. It is therefore more informative to test different delay intervals when comparing WM across different species. Inoue and Matsuzawa only controlled for the exposure time to the numbers to be encoded, rather than the delay between masking and recall. Even so, the claim that the performance of Ayumu is not matched by humans is not correct.

Claims like this deriving from this experiment have been extensively repeated in the comparative psychology and human evolution literature (Balter, 2010). The claim that the performance by Ayumu cannot be matched by humans has been countered by experimental evidence. Cook and Wilson (2010; see also Silberberg and Kearns, 2009) have shown that humans who have followed the same training protocol as did Ayumu perform correctly 90 % of the time, which is statistically greater than the 80 % correct performance of Ayumu. Similarly, Potter (2012) has shown that humans can distinguish a particular photograph image from a sequence of 12 photographs even when each photo is visible for only 1 ms. Potter and colleagues account for this ability by referring to what they call conceptual working memory, which is distinct from working memory as defined by Baddeley (Potter et al., 2004). It should also be mentioned that the effectiveness of the experiment has also been called into question (Cook and Wilson, 2010) due to the way the numbers were masked, which may not have controlled for iconic memory and retinal after-images (Humphrey, 2012).

Altogether, there is neither any evidence showing that Ayumu's performance is a measure of the size of HWM nor that Ayumu has working memory capacities not matched or exceeded by humans. For this reason, we find it puzzling that Ayumu's achievement is held out as providing strong support for general claims about differences in working-memory capacity between chimpanzees and humans.

## 6.6. Chimpanzee working memory capacity is comparable to 3.5-year-old children

### 6.6.1. Experimental data do not support claim for remarkable human-like working memory updating abilities

A recent experiment conducted by Völter and colleagues (Völter et al., 2019) has been widely quoted in the popular media due to its claim that "the current study provides evidence for remarkable WM updating abilities in chimpanzees" (p. 8). Each trial of the experiment involves from 2 to 6 closed boxes that are presented on a platform to a subject chimpanzee. All the boxes are baited initially and the goal for the subject chimpanzee is to open them, one at a time, to retrieve the bait from the box as a reward, but without re-opening a previously opened, hence now empty, box. The measure of interest is the maximum number of boxes selected in a trial before selecting a box without bait. There are three different conditions for the boxes in the experiment: (1) all boxes have different features and are shuffled after each choice so that their exterior appearance provides the key to avoid re-visiting a box (Feature Only condition); (2) all boxes are identical and they are not shuffled after each choice, so the position is informative as to a previously visited box (Space Only condition); and (3) a subject can rely on both the features of the boxes and their relative position to avoid revisiting an empty box (Feature + Space condition). The experiment begins for a subject

with a trial having 2 boxes. On each trial, the subject makes as many choices as there are boxes in that trial. A trial is non-redundant if the subject opens all of the boxes in that trial without opening an empty box. If a subject succeeds in doing several consecutive trials without redundancy (5 for 2 boxes, 3 for 3 boxes, 2 for 4, 5 or 6 boxes), another box is added, the boxes are baited, and the experiment continues with that subject.

For the purposes of comparing the performance of chimpanzees with humans, the measure of interest is the number of boxes selected in a trial before selecting a box without a reward. Chimpanzee performance was compared to human performance for the same experiment, using the same test protocol, carried out by Diamond and colleagues (Diamond et al., 1997) with children aged 3.5–7.0.<sup>4</sup>

Völter et al. (2019) compared only the chimpanzees who reached the 6-box level to the performance of children since the Diamond et al. project always used 6 boxes in a trial. For the Feature Only condition, the single chimpanzee that reached the 6-box level selected an average of 5.67 boxes with rewards before making a mistake. For the Feature + Space condition, the 6 chimpanzees that reached the 6-box level selected 5.62 boxes on average before making a mistake. Völter et al. observe that in the Diamond et al. study, 7-year-old children selected an average of 4.0 boxes in the Feature Only condition before making a mistake versus 5.67 for the single chimpanzee that reached the 6-box level in the Feature Only condition. For the Feature + Space condition, 7-year-old children selected 5.3 boxes on average before making a mistake versus an average of 5.62 boxes for the six chimpanzees that reached the 6-box level. Based on this comparison, they concluded that "the current study provides evidence for remarkable WM updating abilities in chimpanzees. ... Chimpanzees exhibited performance levels comparable with human school-age children in similar self-ordered-search tasks" (p. 8). However, we find that the data do not support their conclusion.

### 6.6.2. Völter et al. interpretation of experimental data is invalid for two reasons

We will now show that their claim is problematic for two main reasons: (1) Diamond et al. used means computed over a random sample of children whereas Völter et al. compared the highest performing chimpanzees to the mean performance of all children and (2) the Diamond et al. data do not support the Völter et al. claim that the chimpanzee performance is comparable to the performance of 7-year-old children. We conclude, instead, that the chimpanzee experimental performance does not support Völter et al.'s claim of the chimpanzee HWM capacity being comparable to that of humans.

**6.6.2.1. Problem 1: highest performers from Pan compared to a random sample of children.** The performance by the chimpanzees is shown in Table 2 for each of the three experiment conditions. Völter and colleagues restrict their comparison of *Pan* to *Homo* to the highest performing chimpanzees, namely the chimpanzees that succeeded in reaching trials with 6 boxes. For the Feature only condition, only the chimpanzee Kofi reached this level, but he is an outlier both for Score and # of Boxes. However, it is the mean performance of the chimpanzees and not the performance of an outlier chimpanzee that needs to be compared to the mean performance by the children, so for this reason the comparison with Kofi is not valid and Kofi will be deleted from the statistical analysis discussed here.

Rather than compare the highest performing chimpanzees to the mean performance of a random sample of children, the comparison

<sup>4</sup> One of the anonymous reviewers commented that, though it is not mentioned in the Völter et al. article, Völter and colleagues rescaled the measure used by Diamond and colleagues so it would match their measure. Whereas Völter and colleagues counted the number of boxes selected before opening an empty box, Diamond and colleagues included the empty box in their count. For this reason, we have reduced the means for the Diamond et al. counts by 1 unit.

**Table 2**  
Chimpanzee Performance Scores.

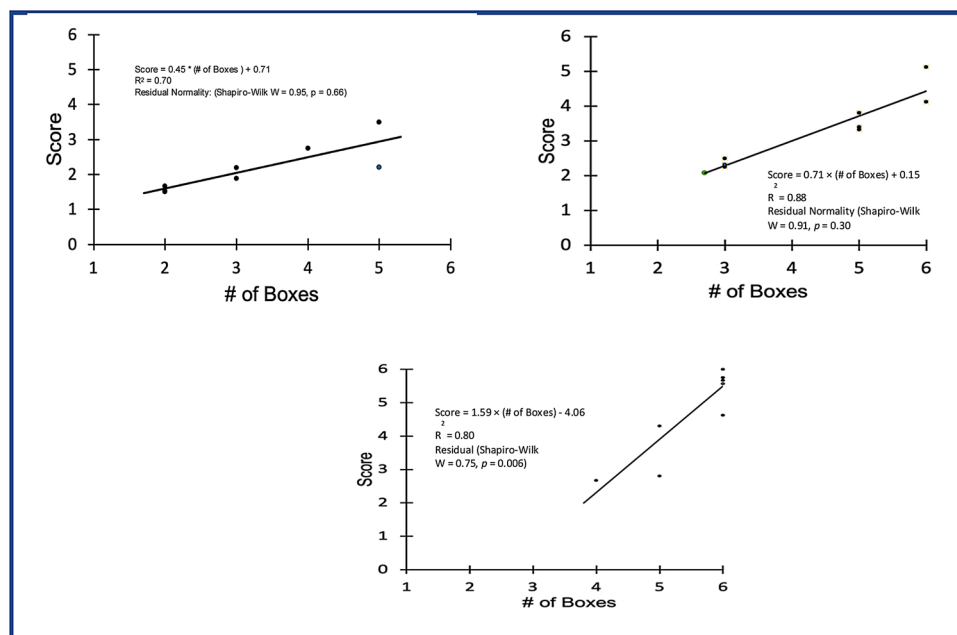
Chimpanzee	Feature Only			Space Only <sup>2</sup>			Feature + Space		
	Score <sup>3</sup>	# of Boxes	n	Score <sup>3</sup>	# of Boxes	n	Score <sup>3</sup>	# of Boxes	n
Dorien	1.67	2	24	3.33	5	12	4.30	5	10
Fraukje	2.20	5	10	2.31	3	16	5.57	6 <sup>1</sup>	7
Frodo	2.19	3	16	3.40	5	10	2.80	5	10
Kofi <sup>5</sup>	5.67	6	6	4.12	6	8	5.67	6 <sup>1</sup>	6
Lobo	1.89	3	9	2.50	3	16	6.00	6 <sup>1</sup>	2
Lome	2.67	4	12	3.80	5	10	5.75	6 <sup>1</sup>	4
Riet	1.50	2	24	5.12	6	8	4.62	6 <sup>1</sup>	8
Robert	1.54	2	24	2.25	3	16	2.67	4	12
Sandra	3.50	5	10	3.80	5	10	5.75 <sup>3</sup>	6 <sup>1</sup>	4
Mean <sup>4</sup>	3.36 <sup>3</sup>			4.43 <sup>3</sup>			5.50 <sup>3</sup>		

- Notes:
1. For the Feature + Space data, Völter et al. (2019) state that only 5 chimpanzees reached the 6-box level, but the Supplementary Material show that 6 chimpanzees reached the 6-box level.
  2. Völter et al. (2019) do not compare the chimpanzees to the children for the Space data since comparable data are not reported in Diamond et al. (1997).
  3. When # of Boxes < 6, Score is extrapolated linearly to 6 boxes for computing the mean score by the equation, Score<sub>extrapolated</sub> = Score<sub>measured</sub> + b × (6 - # of Boxes), where b is the slope for the linear regression of Score on Age, b = 0.44 for Feature Only, b = 0.71 for Space Only, and b = 1.59 for Feature + Space (see Fig. 10).
  4. Means based on extrapolated Score value.
  5. Kofi is an outlier for the Feature Only condition. His score is not included in statistical calculations for the Feature Only condition.
- Data are from Völter et al. (2019: Supplementary Material).

should be between the mean Score for the complete sample of chimpanzees (excluding Kofi in the Feature Only condition) and the mean Score by Age for the children in the Diamond et al. data set. Since the Diamond et al. data are based on trials with 6 boxes regardless of the performance level of the children and there are no chimpanzees other than Kofi that reached the 6-box level in the Feature Only condition, it is necessary to extrapolate the Score values to 6 boxes in the Feature Only condition. The extrapolation will also be applied to the Score values for the other two conditions to maximize the sample size.

The rationale for the extrapolation stems from the fact that when the number *n* of boxes in a trial is increased to *n* + 1, the probability *p<sub>k</sub>* of selecting an empty box on the *k*th box choice, *k* > 1, decreases, all other things being equal. For example, if for each choice of boxes, a box is selected randomly from the set of *n* boxes in a trial, then *p<sub>2</sub>* = 1/3 when *k* = 2 and *n* = 3, but *p<sub>2</sub>* = 1/4 when the number of boxes in a trial is increased to *n* = 4 boxes. (If the strategy for selecting boxes changes, then these probabilities will also change.) Thus, the Score for a

chimpanzee will be greater on average for its *k*th choice when there are more boxes in a trial than when there are fewer boxes. To compare the mean performance of chimpanzees on trials with different numbers of boxes to the mean performance of the children by age, all of whom had 6 boxes in each trial, the effect of the number of boxes in a trial on the performance of chimpanzees should be corrected for directly, but this would require knowing the strategy with which each chimpanzee selected boxes when doing the experiment. Instead, we can extrapolate from the Score and number of boxes in a chimpanzee trial to a predicted score for that chimpanzee if that chimpanzee were selecting boxes in a trial with 6 boxes by fitting a regression line to the scattergram plot of Score versus # of Boxes and then predicting the Score for 6 boxes from the regression line. This will be an overprediction since the experiment increases the number of boxes in a trial when a chimpanzee does the experiment successfully using the current number of boxes in a trial. Thus, the chimpanzees with more boxes in a trial may also be the better performing chimpanzees and so the regression line measures both the



**Fig. 9.** Graphs showing the increase in Score values with # of Boxes in a trial for the Feature Only condition (upper left), Space Only condition (upper right) and Feature + Space conditions (lower center). In all three conditions, the increase is linear; The upper two regression lines satisfy the goodness-of-fit criterion of normally distributed residuals. The lower regression line has residuals with a long-left tail; hence the residuals are not normally distributed, but it may be seen visually that the linear regression line fits the sample data. The increasing regression line slopes from upper left to lower center reflect increase in performance level by the chimpanzees across the three conditions.

effect of the number of boxes in a trial on a chimpanzee's Score and the fact that some chimpanzees may perform better on this experiment than do other chimpanzees.

The regression lines used for the extrapolations are shown in Fig. 9. The extrapolations are executed as follows. Consider the Feature Only condition. The slope  $b = 0.44$  for the linear regression indicates that Score increases by 0.44 units for each box added to a trial. Consequently, the Score value extrapolated to 6 boxes will be:  $\text{Score}_{\text{extrapolated}} = \text{Score}_{\text{measured}} + 0.44 \times (6 - \# \text{ of Boxes})$ ,

**6.6.2.2. Problem 2: diamond et al. data do not support equating chimpanzees with 7-year-old children.** To compare the Pan data with the Diamond et al. data, we first graph the age versus mean performance scores for the children for both the Feature Only and the Feature + Space versions of the experiment conducted by Diamond and colleagues and then compare: (1) the mean extrapolated Feature Only score of 3.36 for the chimpanzees from Table 2 to the Feature Only age trend for children and (2) the mean extrapolated Feature + Space score of 5.50 to the Feature + Space age trend for children (see Fig. 10). For the Feature Only condition, the mean performance of children decreases from 3.5 to 5 years of age and then increases from 5 to 7 years of age (see Fig. 10 [left]). The scores from the Feature + Space experiment only increase with children in the age range from 3.5 to 5 years. The trend for scores from 5 to 7 years of age is flat (see Fig. 10 [right]).

For the Feature Only condition, only the chimpanzee Kofi reached the 6-box level and so he is the only chimpanzee compared to the performance of the children in the Feature Only condition. While his score of 5.67 is comparable to the performance level of 7-year-old children in the Feature Only condition, it is also comparable with the scores for children from 5- to 7-years in age (see Fig. 10, right graph). Even more, as noted above, Kofi is an outlier and so his score is not representative of chimpanzee performance in the Feature Only condition. Thus, his score of 5.67 does not show concordance of chimpanzee performance with the performance of 7-year-old children as claimed by Völter et al. Instead, it is the chimpanzee mean score of 3.36 based on Score values extrapolated to 6 boxes that should be compared with the trend for the children.

This mean score intersects the pattern for the children means at 3.9 years of age (see dashed lines in Fig. 10 [left]) and is consistent with the trend for the children means from 3.5 to 5 years of age (see Fig. 10 [left]). Hence the chimpanzee average scores are consistent with 3.9-year-old children. From Fig. 1B, these data imply  $\text{HWM} = 2.1$ .

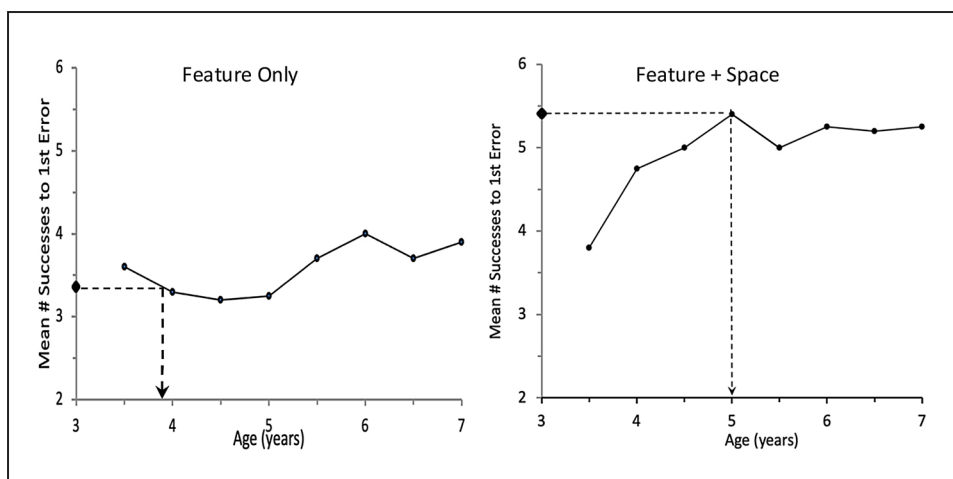
For the Feature + Space condition, while the mean performance of the chimpanzees (whether the mean of 5.62 for the 6 chimpanzees with # of Boxes = 6 or the mean score of 5.50 for the extrapolated score) is consistent with the performance of the 7-year-old children, the performance of the children remains constant from age 5 to 7, as noted above.

Völter et al., do not justify using the right extreme of this range of possible values to characterize the chimpanzees. That the Feature Only data consistently match the chimpanzees with 3.9-year-old children implies the mean of 5.50 for the Feature + Space scores should be matched, at most, with 5-year-old children. This datum implies  $\text{HWM} = 2.7$  (see Fig. 1B).

The range from 2.1 to 2.7 for the size of HWM derived from comparison of chimpanzee performance with the performance level of the children doing the same experiment implies the holding capacity of chimpanzee WM is  $2 \pm 1$ . We conclude, then, that these data neither support the claim by Völter and colleagues that the HWM of chimpanzees is comparable to that of 7-year-old children nor that the chimpanzees have "remarkable WM abilities." Altogether, the Völter et al. data, the Inoue and Matsuzawa data, and the Kawai and Matsuzawa data all support a characterization of chimpanzee HWM as being  $2 \pm 1$ .

**6.6.2.3. Data from chimpanzee experiment are consistent with  $\text{WM} = 2 \pm 1$  in Pan.** We begin this section by noting that the chimpanzees performed better in the space condition than the feature condition. Firstly, the mean performance scores (see Table 2) increase from mean = 3.36 for feature-only condition to mean = 4.43 for the space-only condition to mean = 5.50 for the feature + space condition. Secondly, it is not just that chimpanzees perform better in the space-only condition as a group (i.e., on average), but do so individually as well. In fact, 8 of 9 chimpanzees performed better in the space-only condition than in the feature-only condition. And thirdly, for the feature-only condition, the average highest box level achieved is 3.6 boxes, whereas for the space-only condition the average highest box level achieved is 4.6 boxes.

These results stem, we suggest, from the strategies subjects employed to solve the space-only condition. Table 3 presents a detailed account of the subjects' performance in the space-only condition and shows that irrespective of whether there are 3, 4, 5 or 6 boxes, most of the choices on the first two trials are the two end boxes. On the first trial there are 114 end box choices versus 33 middle box choices and on the second trial there are 112 end box choices versus 35 middle box choices, with a subject's first choice being typically the rightmost box (94 right choices versus 20 left choices) while the second choice is usually the leftmost box (33 right choices versus 79 left choices). These data show that the chimpanzees seem to rely on a simple rule for their first two choices: choose the rightmost box in the initial trial and the leftmost box in the next trial. Since what is rightmost or leftmost stays the same from one trial to the next throughout the experimental session, we hypothesize that this is less demanding of working memory than what is required for the feature-only condition since WM would not need to be brought into play for the first two trials in the space-only condition when using this hypothesized rule. Support for this hypothesis can be found in studies



**Fig. 10.** Age-based pattern for Feature Only and Feature + Space conditions showing the age-based mean number of boxes opened correctly by children (standard deviations not provided) in trials with 6 boxes (Diamond et al., 1997). The diamond on the vertical axis is the mean score for the chimpanzee subjects, with scores extrapolated to 6 boxes when a trial had < 6 boxes (see text for details). In the Feature Only condition, the performance of the chimpanzees matches the performance of 3.9-year-old children and their performance in the Feature + Space condition matches that of 5.0-year-old children. Children data are derived from Diamond et al. (1997: Figs. 21 and 22).

**Table 3**  
1<sup>st</sup> and 2<sup>nd</sup> Choice of Boxes in Hidden Food Box Experiment.

	3 boxes						4 boxes					
	1 <sup>st</sup> Choice			2 <sup>nd</sup> Choice			1 <sup>st</sup> Choice			2 <sup>nd</sup> Choice		
	left	right	center	left	right	center	left	right	center	left	right	center
	1	14	1	15	1	0	0	2	0	0	0	2
	4	8	4	8	6	2	0	2	0	2	0	0
	2	3	0	2	1	2	0	3	3	3	1	2
	0	1	2	0	2	1	0	2	0	0	0	2
	3	11	2	8	6	2						
	2	4	0	4	2	0						
	1	2	0	1	1	1						
	0	4	0	3	0	1						
	0	5	3	5	2	1						
Total	13	52	12	46	21	10	0	9	3	5	1	6
	<b>Either end: 65</b>		12	<b>Either end: 67</b>		10	<b>Either end: 9</b>		3	<b>Either end: 6</b>		6
Expected	51.3		25.7	51.3		25.7	6.0		6.0	6.0		6.0
	$X^2 = 10.96, p < 0.0001$			$X^2 = 14.40, p < 0.0004$			$X^2 = 3.00, p < 0.10$			$X^2 = 0.00, p < 1.0$		

	5 boxes						6 boxes					
	1 <sup>st</sup> Choice			2 <sup>nd</sup> Choice			1 <sup>st</sup> Choice			2 <sup>nd</sup> Choice		
	left	right	center	left	right	center	left	right	center	left	right	center
	3	5	4	5	3	4	0	7	1	2	0	6
	0	5	1	0	0	6	0	8	0	8	0	0
	1	8	0	7	1	1						
	3	3	4	6	1	3						
	0	3	7	1	6	3						
	1	7	2	4	2	4						
Total	8	31	18	23	13	21	0	15	1	10	0	6
	<b>Either end: 39</b>		18	<b>Either end: 36</b>		21	<b>Either end: 15</b>		1	<b>Either end: 10</b>		6
Expected	22.8		34.2	22.8		34.2	5.3		10.7	5.3		10.7
	$X^2 = 18.81, p < 0.0000$			$X^2 = 9.4, p < 0.003$			$X^2 = 26.54, p < 0.0000$			$X^2 = 6.23, p < 0.05$		

Note: Center refers to all boxes between the two end boxes.

that employed a change detection task to investigate working memory and discovered that when target templates do not change across trials, the demand on working memory is low because the template is stored in long-term memory (Woodman et al., 2007). It follows that the space-only condition would only partially test spatial memory defined as the capacity to locate and keep in memory objects in 3-dimensional space if the first two trials are guided by the hypothesized simple, uni-dimensional mechanical rule rather than by remembering spatial locations in 3-space.

This raises the question, though: When there are three boxes, do the chimpanzees preferentially select the middle box, the only one that will have food after selecting the two endpoints, as their third choice? The answer is no. Of the  $n = 57$  times when there are three boxes and the two end points are selected in the first two trials, then on the third trial an end box is selected  $o_e = 34$  times and a middle box is selected  $o_m = 23$  times. For the null hypothesis that the third choice is random, the expected number of end boxes selected on the third trial is  $e_e = 38.0$  and the expected number of middle boxes selected on the third trial is  $e_m = 19.0$ . A chi-square test of the null hypothesis yields  $X^2 = 1.69, df = 1$ , and  $p > 0.15$ , hence the null hypothesis fails to be rejected. This suggests that the first two choices do not involve working memory, so are not loaded into working memory, thus the choice of the end boxes in trials 1 and 2 would not affect the third choice, hence the third choice should be random, as is shown by the data. Given this, WM would be initiated with the 3<sup>rd</sup> and 4<sup>th</sup> trials and so if WM capacity = 2, the 5<sup>th</sup> trial should not repeat the 4<sup>th</sup> trial. This is the case. The 5<sup>th</sup> trial is the same as the 4<sup>th</sup> trial only 7 times out of 76 instances. Next, if the 5<sup>th</sup> trial is then loaded into WM, the same pattern should hold when there are 6 boxes, namely the 6<sup>th</sup> trial should not repeat the 5<sup>th</sup> trial. However, with 6 boxes, the 6<sup>th</sup> trial is the same as the 5<sup>th</sup> trial 6 times out of 16. A chi-square test for comparing the ratio of 7/76 versus 6/16 yields  $X^2 = 8.55, df = 1$ , and  $p < 0.005$ , hence the null hypothesis that the two ratios relate to the previous trial in the same way may be rejected. This indicates that the 5<sup>th</sup> trial is not, or cannot be, loaded into WM, which is also consistent

with the argument that the WM capacity =  $2 \pm 1$  in the chimpanzees.

Next, we compare their performance in the space-only condition with the feature-only condition. Though the feature-only condition includes randomization of the boxes after each trial, the Supplementary Data in Völter et al. (2019) includes both the space and the feature values for the box choice made on each trial. On the first trial and when there are three boxes, the leftmost box is selected  $n_l = 11$  times, the rightmost box is selected  $n_r = 37$  times, and the middle box is selected  $n_m = 12$  times. When compared with the first trial data for the space-only condition from Table 3,  $X^2 = 0.60, df = 2$ , and  $p > 0.70$ , showing that the chimpanzees are using identical strategies for the first trial in both the space-only and the feature-only conditions. For the 2nd trial, the leftmost box is selected  $n_l = 24$  times, the rightmost  $n_r = 19$  times and the middle box  $n_m = 17$  times. When compared with the second trial data for the space-only condition from Table 3,  $X^2 = 7.54, df = 2$ , and  $p < 0.05$ , showing that the performance on the 2<sup>nd</sup> trial in the feature-only condition is a modified form of the performance on the 2<sup>nd</sup> trial in the space-only condition with the leftmost box selected less often than expected based on their performance on the 2<sup>nd</sup> trial in the space-only condition. This reduction is not surprising since the chimpanzees are presumably keeping track of the feature of the box that was selected in the first trial and the box with that feature may, by chance, be in the leftmost position for the 2<sup>nd</sup> trial and so avoided as a choice on the 2<sup>nd</sup> trial.

When there are 5 boxes in the feature-only condition, and if WM capacity =  $2 \pm 1$ , then the performance on the 4<sup>th</sup> trial will not have been uploaded to WM and so comparison of the 5<sup>th</sup> trial to the 4<sup>th</sup> trial in the feature-only condition should be comparable to comparison of the 6<sup>th</sup> trial to the 5<sup>th</sup> trial in the space-only condition since, as discussed above, the 5<sup>th</sup> trial in the space-only condition would not be loaded into WM when WM capacity =  $2 \pm 1$ . In the feature-only condition and with 5 boxes, the trait choice on the 5<sup>th</sup> trial is the same as the trait choice on the 4<sup>th</sup> trial on 8 out of 27 instances. When compared with 6 out of 16 instances having the same space choice on the 6<sup>th</sup> trial as on the 5<sup>th</sup> trial

in the space-only condition,  $\chi^2 = 0.29$ ,  $df = 1$ , and  $p > 0.55$ , hence the two ratios are statistically the same, thus showing that in the feature-only condition the 4<sup>th</sup> trial has not been initialized in WM, hence these results are consistent with WM capacity =  $2 \pm 1$ . Altogether, these data are consistent with WM capacity =  $2 \pm 1$  for the chimpanzees and do not show that chimpanzees have WM capacity comparable to, let alone exceeding, that of humans.

### 6.7. Experimental evidence shows pure working memory size in *Pan* is less than pure working memory size in *Homo sapiens*

#### 6.7.1. Pure working memory and effective working memory defined

All the experiments discussed so far have addressed the performative aspect of the size of working memory. However, Cowan (2001, 2010) makes a distinction between (1) storage-specific WM, whose size he refers to as pure working memory size (PWMS) and (2) process relating short term WM, whose size will be referred to here as effective working memory size (EWMS). The pure working memory size and the effective working memory size, with the former referring to the architecture of working memory and the latter to the implementation of that architecture, are both valid concepts (Cowan, 2001), but differ in that measurement of the former requires eliminating any performance implementation, such as chunking or rehearsal, that can increase the number of units recallable from working memory. Cowan argues that the average size of PWMS =  $4 \pm 1$  in humans with individual measures ranging from 2 to 6.

**6.7.1.1. Distinction between pure working memory and effective working memory.** A recent article – Carruthers (2013) – referring to Cowan's concepts regarding working memory assumes that the size of the holding capacity of pure working memory is the only valid measure for the size of the short-term memory component (i.e., the size of HWM) of working memory, but this leads to confounding the architecture of working memory, which is what pure WM refers to, with the implementation of that architecture, which is what effective WM refers to. The difference between the two can be seen in the fact that, architecturally, a Turing machine (Turing, 1936) has a working memory of size 1, but can, in principle, be programmed to solve any computational problem. Thus, the level of performance of a computational system such as working memory is not determined by its architecture, *per se*, but also by how that architecture is implemented.

Differences in implementation motivates Cowan's comments that PWMS, as a measure of the size of short-term working memory, is not a replacement of EWMS:

... the somewhat higher limit of  $7 \pm 2$  stimuli ... is valid ... as a commonly observed, compound STM [short term memory] limit for materials that allow online rehearsal, chunking, and memorization.... (2001:114)

PWMS, then, refers to the measured size of short-term WM only when information processing that affects the measure of the size of WM, such as verbal rehearsal of items or mental processing that interferes with the storing of information, is excluded (Cowan, 2010). Thus, the experiment reported on by Völter et al. (2019) is not a test of PWMS but of EWMS since the experimental subjects, whether human or chimpanzee, are free to develop strategies for remembering either the space location of the boxes involved in the experiment or the color feature of those boxes.

**6.7.1.2. Notation for distinguishing holding working memory size in *Pan* from holding working memory size in *Homo sapiens*.** To distinguish whether we are referring to the working memory of chimpanzees or to the working memory of humans, we will refer to the pure working memory size of *Pan* by PWMS\* and the effective working memory size of *Pan* by EWMS\*. For *Homo sapiens*, PWMS =  $4 \pm 1$  (Cowan, 2001) and EWMS =  $7 \pm 2$  (Miller, 1956). What we will now show, contrary to the

claim that PWMS\* = PWMS (Carruthers, 2013), is that PWMS\* < PWMS and EWMS\* =  $2 \pm 1$  whereas EWMS =  $7 \pm 2$ .

#### 6.7.2. Experimental evidence for pure working memory size in *Pan*

Carruthers makes the claim that PWMS is comparable, in general, for animals: "... experimental work with animals suggests that their [pure] WM limits may fall within the human range" (2013, p. 10373). If this claim were valid, then showing PWMS\* = PWMS could proceed indirectly by showing that PWMS in monkeys is the same as in humans. He cites two experiments in support of this claim. The first experiment, conducted by Botvinick et al. (2009) using a macaque monkey, does not need discussion here since Carruthers states that the experiment does not establish that PWMS = 4 in the macaques, as the experiment is "consistent with a claimed WM limit of one or two items" (2013: 10373). The only evidence he provides, then, in support of his claim that PWMS = PWMS\* is an experiment reported on by Hauser et al. (2000).

**6.7.2.1. Experiment measures subject's ability to distinguish among small quantities through subitizing.** The experiment involves putting a small number of slices of apple in one opaque box and a larger number in another box while the monkey is watching. Hauser et al. report that the monkey will then choose the box with a larger number of food items for two versus three items, three versus four, or three versus five items but not for four versus five, four versus six, four versus eight or three versus eight. Since the monkey can, apparently, distinguish up to four items, Carruthers interprets the result of this experiment as showing that PWMS = 4 for monkeys.

However, as Beran and Beran (2004, p. 95) comment, the performance of the monkeys is consistent with "a limited-capacity system of representation" such as "the object file system" that "can represent only up to four items in a set because of the visual system's limits for the parallel individuation of items in a presented set (Trick and Pylyshyn, 1994)." The object file representation system has been used to model subitizing, the term used to refer to recognition of up to four items in a collection without counting the items, a recognition process that occurs in humans and many other animals. The object file representation system allows for the comparison of one set of objects with another set so long as there are at most four objects in each set, hence the failure of the Rhesus (macaque) monkeys to distinguish between four versus five, four versus six, four versus eight or three versus eight slices.

**6.7.2.2. Performance level on experiment is not a measure of pure working memory.** Carruthers assumes, but does not establish, that the Hauser experiment measures pure working memory. However, experiments with other apes using the same research design show a significant problem with his assumption. Beran and Beran (2004) conducted the same experiment with a sample of four chimpanzees. Each chimpanzee was able not only to distinguish between comparisons with  $\leq 4$  bananas placed in each container, thus matching the performance of the Rhesus monkeys in the Hauser et al. experiment, but was equally able to distinguish between 5 versus 8, 5 versus 10 and 6 versus 10 bananas. If, as claimed by Carruthers, the experiment is a measure of pure working memory, then PWMS\* = 10 for the chimpanzees, more than double the value for humans, which obviously is not the case.

The Beran and Beran experiment has been extended to include a larger sample of chimpanzees as well as a sample of gorillas and orangutans (Hanus and Call, 2007). The results with the larger sample are like those from the Beran and Beran (2004) experiment. Regardless of species, each ape was able to distinguish between smaller and larger numbers of pieces of food in the two opaque containers, for up to 10 pieces of food. Hanus and Call suggest that quantity estimation is being done with an analog estimator which is not present in Rhesus macaques. Alternatively, possibly both a subitizing enumeration system and an analog enumeration system are present in the apes, but only a subitizing system is operative in the Rhesus monkeys.



6.7.2.3. *Subitizing is distinct from working memory.* This raises the question as to whether subitizing relates to the magnitude of PWMS at all. Though it has been suggested (see [Cutini and Bonato, 2012](#)) that subitizing in humans may relate to visual short term memory since human visual short term memory is of size 4 and humans easily subitize collections up to size 4, this does not account for the performance by the Rhesus monkeys as their visual short term memory is about one-half that of the visual short term memory for humans ([Elmore et al., 2011: Fig. 2A](#)), hence they should only be able to deal with collections of size 1 or 2 if the size of visual short term memory is the basis for their performance. In addition, research done by [Ashkenazi \(2016, p. 5\)](#) shows that the Space component of visual short-term memory does not relate to subitizing:

Subitizing is supported by a domain general visual object individuation mechanism and not a Space working memory mechanism. It is proposed that subitizing is supported by occipital-temporal areas but not the inferior parietal cortex.

Thus, there is no reason to assume that the enumeration performance of the Rhesus monkeys implies their  $PWMS^* = 4$ .

6.7.3. *Growth span for pure working memory differs between Pan and Homo sapiens*

Another difficulty with the claim that  $PWMS = PWMS^*$  for the chimpanzees is that difference in the time spans for the growth in the size of PWMS in humans versus chimpanzees would yield different final values for PWMS. To see this, consider first the time span for growth of PWMS in humans.

The onset of active working memory in human infants starts as early as 5.5 months ([Reznick et al., 2004](#)) with  $PWMS = 2$  for infants of age 5–7 months ([Feigenson, 2007 and references therein; Moher et al., 2012 and references therein](#)), though others (e.g., [Káldy and Leslie, 2003](#)) argue that  $PWMS = 1$  for this age range. By age 10–14 months, PWMS has increased to 3 ([Feigenson and Carey, 2005 and references therein; Moher et al., 2012 and references therein](#)) and PWMS then increases further to its mature working memory size of  $PWMS = 4$ .

While growth in EWMS does not match the pattern for increase in brain size during human development ([Read, 2008 and references therein](#)), the increase in PWMS should, however, relate to growth in brain size since PWMS refers to the architecture of the brain rather than to the implementation of that architecture:

The prefrontal cortex of the brain and the executive functions it supports undergo a long period of development in human ontogenesis, the fastest development occurring in the child's first years of life.... Out [sic] data support the existence of continuity in the development of the mechanisms of working memory associated with activity in the prefrontal cortex of the brain during the first years of the child's life. ([Tsetlin et al., 2012, p. 698](#)).

Since rapid brain growth ceases after age 36 months (see [Fig. 11](#)), we

may (conservatively) assume that the adult value of  $PWMS = 4$  is reached by age 36 months, at most.

6.7.3.1. *Correlation between growth trajectory for pure working memory and increase in brain weight.* Over the time period from 6 months to 36 months there is virtually a perfect correlation between the growth trajectory for PWMS and increase in brain size measured by brain weight for males, with a slightly less strong correlation for females ([Fig. 11](#)). If we extrapolate backwards from 6 months to birth, we get a predicted value of  $PWMS = 1$  at birth for humans with a brain weight about 400 g. The growth trajectory in PWMS implies, then, an increase of 1 unit in PWMS for each 400 g of brain weight. That brain weight does not reach 800 g until age 12 months may account for why some researchers find that  $PWMS = 1$  at age 6 months, while others find that  $PWMS = 2$  at that age (see references above). The average brain weight would be about 600 g at 6 months, hence in a cohort of 6-month-old infants, PWMS would still be in transition from  $PWMS = 1$  to  $PWMS = 2$  and so some infants in this cohort would have  $PWMS = 1$  and other infants would have  $PWMS = 2$ . Thus, the difference in the reported values for PWMS by different research groups may be due, in part, to a sampling effect.

6.7.3.2. *Human growth rate applied to Pan implies  $PWMS^* < PWMS$ .* Next, we can apply these results to the time span and amount of growth in brain weight in chimpanzees but recognizing that a correlation determined from one species need not hold in the same manner in a related species, even in a sister species. This is what we find with humans versus chimpanzees. If we apply the human correlation between brain weight and PWMS directly to chimpanzees without modification, then for their average brain weight at birth of 151 g and average adult brain size of 382 g ([DeSilva and Lesnik, 2006](#)), the human data of 1 unit of PWMS per 400 g of brain weight would imply chimpanzees have  $PWMS^* = 0$  at birth and  $PWMS^* = 1$  at adulthood, which seem to be too small.

Alternatively, if we assume that in humans the main evolutionary change in the size of PWMS has been an increase in its upper bound and not in the starting size of PWMS at birth from  $PWMS^* = 0$  for chimpanzees to  $PWMS = 1$  for humans, then we would posit, by homology between  $WM^*$  and  $WM$ , that  $PWMS^* = 1$  at birth and the growth in brain size of 231 g between birth and chimpanzee adulthood would only add, at most 1 unit of PWMS, implying that  $PWMS^* = 2$ , at most, for adult chimpanzees. Either way, the  $PWMS^*$  for chimpanzees would, at most, be about one-half the PWMS for humans, not the  $PWMS = 4$  claimed by Carruthers for chimpanzees.

For chimpanzees to have the same PWMS as humans, there would have to be a much higher growth rate for  $PWMS^*$  in chimpanzees per unit of added brain mass than is the case for humans, hence there must have been neotenzionization of cognitive development associated with brain growth in humans, but there is no evidence for neotenzionization in human cognitive development ([Langer, 2005, 2006; Parker and McKinney, 1999](#)). Thus, the more plausible assumption is that for adult

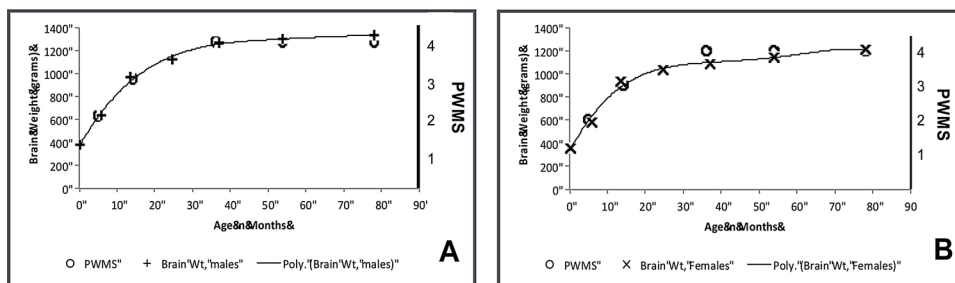


Fig. 11. Change in brain weight (“+”) with age compared to change in PWMS (“o”). PWMS is assumed to reach its maximum by 36 months of age when rapid growth in brain weight ceases. A 4th degree polynomial has been fit to the brain weight data to illustrate the time-trend in brain weight. Data on human brain weight are from [Dekaban \(1978: Tables 2 and 3\)](#). (A) Male data. For the five data points for PWMS, the correlation  $r = 0.974$  between brain weight and PWMS is statistically significant ( $p < 0.01$ ,  $n = 5$ ) at the  $\alpha = 1\%$  level. (B) Female data. For the five data points for PWMS, the correlation  $r = 0.947$  between brain weight and PWMS is sta-

tistically significant ( $p < 0.015$ ,  $n = 5$ ) at the  $\alpha = 5\%$  level.

chimpanzees PWMS\* is around  $2 \pm 1$  in comparison to humans with  $PWMS = 4 \pm 1$ , which implies that limitations on chimpanzee EWMS are not simply due to the various attributes that differ between WM\* and WM as argued by Carruthers, but to chimpanzees having a smaller PWMS to begin with.

## 7. Comparison of SWM and AWM sub-systems of working memory in humans and in chimpanzees

Attention has been directed in the foregoing paragraphs to the *holding* sub-system (HWM) of working memory. It is no less important to consider the *strategic* (SWM) and *attentional* (AWM) sub-systems of working memory (for hierarchical planning) and the computational power of working memory which enables mental manipulation. For this purpose, we consider four sets of experiments involving chimpanzees that relate to different cognitive levels: one relates to planning abilities (SWM), a second to prospective and strategic planning (SWM) capacities and the last two to whether chimpanzees exhibit theory of mind (AWM).

Before assigning a particular cognitive level normally associated with human behavior to chimpanzees collectively, though, it is not sufficient to just identify a single chimpanzee that can seemingly perform at the cognitive level in question. In addition, we need to examine carefully whether the data said to demonstrate individual performance at the cognitive level in question can be generalized to chimpanzees as a collectivity, considering possible methodological problems involving selection of participants, design of experiments, and quantification of data. To focus on the problems, we now consider four studies involving evidence of problem-solving, planning, and theory of mind by chimpanzees, and we find that either the purported cognitive level does not occur or it occurs through individually exceptional performance that does not carry over to chimpanzees as a collectivity. This, in conjunction with the evidence already presented of a substantial difference in the size of WM (HWM to use our previously defined terminology), invalidates claims that chimpanzee WM performance is comparable to that of humans.

### 7.1. Four sets of experiments addressing planning, future planning and theory of mind abilities in Pan

The first two sets of experiments relate to the capacity of the strategic working memory of great apes. The third and fourth experiment sets relate to attentional working memory. We review each of these sets of experiments to see whether an experiment succeeds in demonstrating that the chimpanzees have the advanced cognitive abilities that some have claimed is part of the cognitive repertoire of the chimpanzees and possibly of other great apes.

#### 7.1.1. Experiment set 1: planning capability of Pan (SWM sub-system)

The first experiment we review was conducted by Tecwyn and colleagues twice: first with bonobos and orangutans as subjects (Tecwyn et al., 2013) and second with human children as subjects (Tecwyn et al., 2014). In both cases, the experiment involved an apparatus with 3 tiers of rotatable paddles. By rotating the paddles leftwards, rightwards, or flat, the subject could direct food placed as bait on a paddle at the top tier downward, to the bottom tier along an uninterrupted path by anticipating how the food will drop from one tier to the next. The subject's goal is for the food to arrive at the only box at the bottom of the apparatus from which the reward can be collected.

Two conditions controlled the difficulty of the experiment. Condition 1 required advanced planning. For the subject to succeed, the paddles had to be turned to form a pathway from the top tier to the box at the bottom of the apparatus before the paddle at the top tier with the food is turned. In Condition 2, the paddles were initially set in the flat position and could then be turned sequentially to direct the food to go from one tier to the next in a stepwise manner. Tecwyn et al. (2014) comment, regarding the difficulty of the experiment:

The sequential planning task is visuospatially simpler because all of the paddles are [initially] flat, which makes mentally visualizing the movement of the reward to the goal more straightforward. The item can also be controlled in a step-by-step manner. Therefore, it is possible to use a perceptually guided strategy (turn the paddle with the item on it toward the open goal) that incrementally moves the item toward the goal. ... The advance planning task, on the other hand, is visuospatially more complex because it is not possible to follow a visual path from the reward to the goal given that relevant paddles need to be pre-positioned before the start paddle is rotated in order to create a valid path. (p. 97)

Both the bonobos and the orangutans failed in the advanced planning condition, and indeed in 84.9 % of the trials for orangutans and 98.3 % for bonobos only the start paddle was rotated. They both succeeded, however, in the sequential planning task. This did not require advanced planning.

When the same experiment was replicated with human children, Tecwyn and colleagues divided the children into three age groups: 4 - 5, 6 - 7 and 8 - 9-years-old. The 4 - 5-year-old children performed comparably to the apes under both conditions. The 6 - 7-year-old children outperformed the apes in both conditions, thus showing advanced planning skills. Taken at face-value, these results imply that the size of the ape HWM is  $2 \pm 1$  (see Fig. 1B). The results of these two studies run counter to the opinion derived from field observations of tool-use by apes in the wild that they have remarkable planning capabilities (Brewer and McGrew, 1990; Boesch et al., 2009).

#### 7.1.2. Experiment set 2: planning for future events by Pan and Pongo (AWM sub-system)

The second study was aimed at assessing the extent to which bonobos and orangutans have the human cognitive ability to plan for a future, possible event. This involves prospective memory, which underlies the ability to carry out in the future a possible action or event conceptualized or recognized in the present and is linked to working memory in the following manner.

It has been posited (Lewis-Peacock et al., 2016 and references therein) that there are two routes for prospective memory performance. One involves WM to monitor the environment for the appropriate future time to undertake the event and the other involves episodic memory for storing the intention to enact the event and then activating the stored intention at the right time. That is, working memory relates to proactive control (Lewis-Peacock et al., 2016) through WM executive functions that serve "(a) to coordinate performance on multiple tasks, (b) to attend to some and inhibit other information, and (c) to hold and manipulate information in long-term memory" (Kliegel et al., 2002, p. 304), whereas episodic memory relates to reactive control of implementing events in the future (Lewis-Peacock et al., 2016). It is the connection with working memory that is of concern here. In this regard, working memory executive processes and prospective memory performance are closely related (Kliegel et al., 2002 and references therein), with > 44 % of the variance in prospective memory performance accounted for by working memory executive functions (Kliegel et al., 2002). We can relate, then, prospective performance of the ape subjects in this study to their WM capacity, albeit with the caveat that other factors can be involved as well.

The future event for the ape subjects was formulated in the following manner. Each of the 3 orangutan and 3 bonobo subjects in the study was offered 8 tool objects of which two alone were appropriate for attaining a food reward in sight, albeit blocked inside a transparent apparatus (Mulcahy and Call, 2006). Each subject was then taken into another room, could take tools with him or her and after a fixed period (1 or 14 h depending on the trial), was then returned to the food apparatus from which any tools not taken by the subject had already been removed. The block was released and if the subject had taken a functional tool, it could now be used to get a food reward. Two of the three bonobos chose more

functional tools than might have been expected by chance, seemingly implying a capacity to plan for an uncertain but possible future need; however, one of the bonobos performed no better than chance (Mulcahy and Call, 2006: Table 1). In addition, it is not clear that the food apparatus was blocked from the subjects' view while they waited in the adjacent cage during the 1-hour waiting interval (there is no information presented on this matter in the Materials and Methods section of the article). Thus, a subject might still be aware of a food reward throughout each trial of the experiment, which obviates any need to plan for a future, uncertain event. This would not apply to the 14-hour waiting interval when subjects were moved upstairs to their sleeping rooms before testing was resumed.

In a related experiment designed to inquire whether chimpanzee, bonobo or orangutan subjects make tools for future use (Bräuer and Call, 2015), a frangible wooden board had to be broken into smaller pieces that could be used to dislodge grapes from 8 transparent tubes mounted in a battery beside their cage baited with 8, 3, or no grapes, with 1 grape per tube. The subjects could prepare tools before the tubes were unblocked, thereby making the grapes inside the tubes accessible. The experimenters found that the subjects, rather than making tools for possible future use, made most tools after the tubes were unblocked. The only evidence for anticipatory behavior is that rarely were more than two tools made when 3 grapes or none were in the tubes, though for dislodging 8 grapes it was enough to break up the board into smaller pieces (an act that itself could have been a distractor from a visible but unattainable reward; cf., Peake et al., 2002). Moreover, when the access time to the grapes was reduced and extra pressure for preparing tools in anticipation (e.g., before the apparatus was unblocked) was imposed there was no increase in the number of tools produced. The study, at best, corroborates the limited capacity for prospective and strategic planning capacities shown by chimpanzees and orangutans in the previous experiment. As discussed above, one explanation for the limited prospective planning performance by the ape subjects would be a limited working memory capacity.

#### 7.1.3. Experiment set 3: theory of mind in Pan (AWM sub-system)

Another set of experiments used the so-called "goggles" experiment (Heyes, 1998) in which animals are to be trained with goggles with one rim color corresponding to transparent goggles and another rim color to opaque goggles. The aim is to find out whether, based on its own experience, an animal might attribute mental states to others that appear to have the same visual experience. The hypothesis being tested is: Do animals attribute mental states to others that appear to be experiencing the same conditions that they have experienced? The experiment (Karg et al., 2015) was carried out by training the apes with one color and transparency of goggle-like masks on every other day and the other color and opacity of goggle-like masks on the alternate days. Once they were familiar with the properties of the masks, a researcher put on one of the two masks and then turned his head to see whether they followed his gaze selectively as a function of the transparency of the mask. This was not the case and therefore chimpanzees did not seem to be using their own visual experience to infer what another can see. In contrast, human infants at 18 months followed the gaze of a blind-folded experimenter more often when they had experienced see-through blindfolds than after wearing opaque blindfolds (Meltzoff and Brooks, 2008).

Another study on gaze-following found that chimpanzees and bonobos followed an experimenter's gaze more frequently when directed toward a barrier containing a translucent rather than an opaque window (Okamoto-Barth et al., 2007). Contrary to the goggle study, the chimpanzees appeared to consider what the experimenter could or could not see.

A further experiment was conducted to account for the different results that were obtained, using competition over food with a human (cf., Hare, 2001). The chimpanzees were trained with three different lids that were either transparent, opaque, or screened. Screened and opaque

lids looked the same from the outside, though the screened version allowed objects to be seen that were placed below it. A chimpanzee sitting face-to-face with the human had to choose between stealing a peanut from under one of two lids, one of which was always opaque and the other was similar in appearance to the opaque lid but was either transparent or screened, hence the human could see through it. The human removed the peanut whenever the human could see the chimpanzee trying to grab it but could not see the chimpanzee reaching under the opaque lid.

The hypothesis being tested was whether the behavior of the chimpanzee varied as a function of a social condition (a human was present and could steal the peanut) or a non-social one (the chimpanzee was alone). Chimpanzees preferred to steal the peanut from under the opaque lid more often in the social condition than in the non-social condition, which was interpreted to support their capacity for visual perspective taking. Nevertheless, in the social condition the apes still tried to steal food from under the transparent and screened lids in more than 40 % of the trials. This counters the hypothesis since neither any additional cost was required to steal food from under the opaque lid nor was a greater reward obtained from under the transparent or screened lids. Why, therefore, should they risk losing the peanut by trying to steal it from under the transparent and screened lids in a social condition? If, from their own experience, they could infer what the human could or could not see, there should be no transparent or screened lid stealing attempts at all. Therefore, this study implies that, even with a simple theory of mind task, only involving visual perspective attribution rather than more complex knowledge or belief attribution, chimpanzees can adopt the visual perspective of others only with difficulty, possibly owing to limitations in their attentional sub-system of working memory (AWM) that hinders efficient switching from their own perspective to that of another. Or perhaps because, unlike humans, they fail to see others as intentional agents like themselves (Tomasello, 2014). In contrast, the human AWM system allows human subjects to do two different, albeit related, things: 1) inhibit at will their own visual perspective and 2) shift at will from own to the other's visual perspective.

These studies show, at most, that chimpanzees (and possibly the other great apes that were also subjects for these experiments) can, to a degree, solve new problems by inventing new solutions, plan for future needs, and have at least a limited theory of mind. They undeniably have cognitive skills that rely on the power of working memory to store and manipulate mental objects and compute plans, but it is not evident that  $WM > 2 \pm 1$  is necessary. Instead, these studies show the limited capacity of apes to draw up and execute plans (strategic sub-system of working memory) and to switch attention from the perspective of oneself to that of others (attentional sub-system of working memory). The limitations are revealed by low rates of success overall, with relatively few successful individuals observed across the different experiments. We find that while an ever-increasing number of studies attribute human-like cognitive abilities to chimpanzees and other great apes, thereby seeming to suggest that there are no fundamental differences between their working memory and ours, this conclusion disappears both when evidence showing development differences between humans and chimpanzees with the age-based development of working memory is considered and when scrutiny of individual performers is contrasted against analyses of group behavior.

#### 7.1.4. Experiment set 4: false belief and theory of mind in Pan (AWM sub-system)

Whether the great apes, in general, and the chimpanzees, in particular, cognitively recognize that the behavior of conspecifics may be driven by false beliefs they hold about the situation in the world is unclear (Call and Tomasello, 2008). In experiments involving food choice, ape subjects have not demonstrated understanding of false beliefs about the state of the world (Hare et al., 2001; Kaminski et al., 2008; Krachun et al., 2009). Other experiments have been critiqued for not excluding

the possibility that the performance claimed to be indicative of understanding false beliefs could be explained by simple behavioral rules such as “search for things where they last saw them” (Krupenye et al., 2016; Scarf and Ruffman, 2017; Perner and Ruffman, 2005)” (Kano et al., 2019). A recent experiment conducted by Kano and colleagues (Kano et al., 2019) purports to circumvent these criticisms and show that apes understand false beliefs. However, the experimental data do not support this claim.

The experiment involves a subject ape observing the same event under two different conditions involving a barrier: (1) the barrier is opaque or (2) the barrier is translucent. The apes involved in the experiment are divided into two groups corresponding to the two conditions. Each group is first familiarized with the barrier for the experimental condition in which the members of that group will participate as subjects. For both conditions, two boxes are located between the subject ape and an actor who can be observed by the subject ape. The subject ape observes an actor watching a stone being placed into a target box. Then the barrier is placed between the actor and the boxes. Next, the ape sees the stone being removed from the target box and placed into the other (distractor) box before it is removed completely, after which the barrier is then removed. Lastly, the actor moves towards the boxes, staying equally distant from each of them. The eye movement of the ape is tracked to see if the subject’s eye movement is biased towards the target box. Results relevant for our purposes here from the two experimental conditions are presented in Table 4.

The authors claim that under Condition 2 the subject ape attributes “to the actor a true belief that the object had been removed and with having *no expectation that the actor would search in any particular place*” (emphasis added). Thus, the prediction is that a subject ape making eye movement under Condition 2 will glance randomly at the target box and/or at the distractor box. However, this prediction is contradicted by the chimpanzee frequency counts shown in Table 4 for the DLS values computed for eye movement under Condition 2. From Table 4, of the 17 apes that make eye movement under Condition 2, 13 either have DLS = 1 (only glance at the target box) or DLS = -1 (only glance at the distractor box), in contradiction to the prediction, while only 4 apes have DLS in the range [-0.65, 0.50] (i.e., the ape looks at the target and the distractor boxes randomly), hence support the prediction.

Under the second condition where the barrier is translucent, an ape realizes (at best) that if s(he) is at the location of the actor, s(he) would see that the stone was first put into the target box, then it was moved into the distractor box and finally it was removed completely. Under the first condition, with the opaque barrier, the ape would realize (at best) that s(he) can only see where the stone was initially placed, namely into the target box. The eye glances of the apes, under both conditions, can thus be explained by the simple rule: expect the stone to be in the box where

it was previously located. Either the target box or the distractor box equally satisfies the “previously located” criterion for the stone, hence we would expect that, under Condition 2, the proportion of apes glancing towards the target box should match the proportion of apes glancing towards the distractor box. From Table 4, 5 apes glanced at the target box and 8 apes glanced at the distractor box. For the null hypothesis that glancing at the target box or glancing at the distractor box are equally likely, the chi-square value is  $X^2 = 0.69$  with  $df = 1$  and  $p = 0.41$ , so the null hypothesis is not rejected. Contrary to the authors’ claim: “our current results are not explained by behavior rules” (p. 5), the rule, *search where an object was previously located*, accounts for the ape behavior under Conditions 1 and 2 without invoking Theory of Mind (ToM). This rule is cognitively comparable to the behavioral rule, *search where an object was last seen* (though it has a different meaning), that has been used to account for the behavior of the apes in a previous study on false belief attribution (see Krupenye et al., 2016). There is no need, then, to invoke ToM to account for the behavior of the apes in this experiment. A more formal argument reaching the same conclusion is provided by Henley and Povinelli (2020).

Until recently there was consensus that ToM, as measured in children with change of location tasks that require children to verbally report where an agent would search for an object as a function of the agent’s epistemic state (e.g., whether the agent had witnessed the object’s change of location) emerged in early childhood (circa 4). This seemingly protracted maturation of ToM would be related to prior linguistic and/or executive control maturation (e.g., Perner, 1991; Wellman and Cross, 2001). However, recent studies that use different paradigms not relying on language (e.g., anticipatory looking [AL] and violation of expectations [VoE]) have found that 2-year-old children show action understanding that involves considering the agent goals and (and herein lies the rub) the agent’s epistemic status (knowledge and beliefs) (e.g., Onishi and Baillargeon, 2005; Southgate et al., 2007; Surian and Geraci, 2012). This has generated an animated debate among researchers that is partially resolved by proposing two different systems for belief reasoning: one that scarcely uses cognitive resources but is inflexible (i.e., implicit mindreading), and one that is cognitively demanding (i.e., explicit mindreading) since it requires executive control yet remains flexible (Apperly and Butterfill, 2009). The former system would relate to AL paradigms such as the one employed in chimpanzees by Krupenye et al. (2016) and the later to more conventional tasks relying on language (e.g., the Sally-Anne test). Although both systems are specialized for dealing with mental states, the implicit system emerges early on and constitutes a fast and almost automatic way of tracking mental states, while the explicit system puts higher demands on executive functions. According to this distinction, chimpanzees’ success in the “belief attribution” task by Krupenye et al. (2016) might be recruiting the implicit system, in which case working memory would only play a minor role in the performance of subjects. It should be noted, however, that there are problems with replicability in false belief tasks with children using AL as a dependent variable (Kulke and Rakoczy, 2017). It is also noteworthy that the specific design employed by both Krupenye et al. (2016) and Kano et al. (2019) with great apes shows lower replication rates than other designs based on AL (for a comprehensive discussion see Horschler et al., 2020).

## 7.2. Group level versus individual level cognitive abilities

When we speak of human cognition, we establish transitional periods when children’s cognitive abilities undergo qualitative changes (e.g., Piaget, 1959). One benchmark may be the cognitive development in children of the capacity to attribute false beliefs to other minds explicitly. In a sample of 46 Austrian children, none of the 3 – 4-year-old children passed the “Sally-Anne” test, while 57 % of the 4 – 6-year-old and 86 % of the 6 – 9-year-old children passed it (Wimmer and Perner, 1983; Baron-Cohen et al., 1985). Those who passed the test can attribute false beliefs to other minds, and even explain why others have

**Table 4**  
DLS<sup>a</sup> Values for Theory of Mind Experiment.

	Condition 1 $m^b$	Condition 2 $n^b$
Eye Movement <sup>c</sup>		
DLS		
1	7	5
[-0.65, 0.40] <sup>d</sup>	5	4
-1	1	8
No Eye Movement <sup>c</sup>	10	7

Notes.

a. Differential Learning Scale:  $DLS = (\# \text{ target} - \# \text{ distractor}) / (\# \text{ target} + \# \text{ distractor})$ .

b.  $m, n$  – number of chimpanzee subjects.

c. Data values for  $m$  and  $n$  corresponding to *No Eye Movement* are from Kano et al. (2019: Table 1) and data values for  $m$  and  $n$  corresponding to *Eye Movement* were determined from Kano et al. (2019: Fig. 2).

d. Range of DLS values between 1 and -1. Condition 1: Mean DLS = 0.12; Condition 2: Mean DLS = -0.10.

wrong knowledge about the world (i.e., full-blown explicit false belief attribution).

The age at which children shift from not passing to passing the “Sally-Anne” test, it should be noted, is not culturally invariant but affected by both broad cultural differences such as an individualistic versus a collectivist cultural emphasis or by narrower cultural differences such as the age at which formal schooling begins. In one study (Kim et al., 2020), German children (from an individualistic culture) were found to pass the “Sally-Anne” test at an earlier age than Japanese children (from a collectivistic culture). In another study the pattern was shown to be more complex. In a comparison of United Kingdom, Italian and Japanese children, Hughes et al. (2014) characterized the United Kingdom and Italian children as coming from individualistic cultures, the Japanese children from a collectivistic culture, and though the United Kingdom children and the Italian children were both from individualistic cultures, they differed by the fact that the United Kingdom children start formal schooling at an earlier age than the Italian children. The researchers report that the United Kingdom children passed the “Sally-Anne” test at a younger age than either the Japanese or the Italian children, showing a broad cultural effect on performance of the United Kingdom versus Japanese children, and a narrow culture effect with the difference in performance of the United Kingdom children versus the Italian children. At the same time, the Italian children and the Japanese children both pass the test at the same age, thus showing similarity despite the Italian children coming from an individualistic culture and the Japanese children from a collectivist culture. Their similarity in performance suggests that the late schooling of Italian children reverses the individualistic cultural effect that otherwise would lead to a difference between the Japanese and the Italian children in the age when they pass the “Sally-Anne” test. Thus, both broad and narrow cultural differences affect the developmental trajectory of children for false beliefs.

Cultural and schooling effects are not limited to the age at which children comprehend false beliefs. They also affect the level of working memory performance. Several studies have found that illiterates exhibit decreased performance in comparison to literates on working memory tests such as repeating lists of pseudowords and the word pair association test that relate to the phonological loop component of working memory (see, e.g., Castro-Caldas et al., 1998; Gomez-Perez and Ostrosky-Solis, 2006; Kosmidis et al., 2006; and Reis and Castro-Caldas, 1997). This is not surprising because a phonological loop is one component of WM and illiterates score lower than literates on phonological processing. In the other direction, it is also found that formal schooling can enhance memory skills measured by working memory tests such as recalling lists of pseudo words (Kosmidis et al., 2006). Although schooling has been shown to have a positive effect on cognitive abilities in Western societies with their cultural emphasis on schooling, other societies that do not culturally emphasize schooling, such as (extant) hunter-gatherer societies, place cultural emphasis on other kinds of knowledge, such as ecological knowledge, and this may have the same effect on cognitive abilities as does schooling in Western societies (Reyes-Garcia et al., 2016).

### 7.2.1. Experiments show individual but not group level cognitive abilities in Pan

When it is reported that an individual chimpanzee exhibits any of the so-called “uniquely human” cognitive traits, it is not shown that most, if not all, adult chimpanzees tested will exhibit these abilities. In fact, a considerable number of adult subjects never solve some of the tasks for which they are being tested and perform as though they do not possess the trait in question, such as failure to ever crack nuts by some chimpanzees at Bossou, failure to select a tool in anticipation of future need for that tool in the experiment by Mulcahy and Call (2006), failure to perform beyond the cognitive level of 3 or 4 year old humans as shown by data on knowledge and belief attribution (Karg et al., 2015, 2016, Kaminski et al., 2008; Krachun et al., 2010), and so on.

Studies like these also include claims that great apes exhibit extraordinary feats of working memory rivalling those of humans (Inoue and Matsuzawa, 2007; Völter et al., 2019), make tools with foresight (Bräuer and Call, 2015), save them for future use (Mulcahy and Call, 2006), solve food-extraction problems insightfully as if by mental trial and error (Mendes et al., 2007), and adopt human visual perspectives implying a theory of mind (Karg et al., 2015, 2016). When these results are taken at face value, it appears that there are no qualitative differences between human and great ape working memory. Yet in the experiment testing the capacity to save tools for future use, in 9 out of 16 trials 5 of the 6 apes tested failed to return to the test room with a suitable tool after having waited elsewhere for just 1 hour (Mulcahy and Call, 2006), revealing inability to anticipate future needs correctly in > 50 % of attempts. Similarly, apes were inefficient when required to make tools for use in a future extractive task (Bräuer and Call, 2015), being barely able to produce 2 tools beforehand when 8 were needed, usually beginning preparation only after having access to the food but not in anticipation of access. As for theory of mind, chimpanzees failed to infer what a human could or could not see as a function of the mask worn and their own previous experience with the masks (Karg et al., 2015), in marked contrast to the success of 18-month-old children in a similar task (Meltzoff and Brooks, 2008). Even after the testing protocol was modified to be in their favor, the performance of the apes was still only slightly improved.

Our intent here has not been to conduct an exhaustive literature review but to focus on studies that have “allegedly” been a breakthrough in the knowledge we have about prospective/strategic capacities in great apes or their ToM abilities. Our goal has been to show that what at first glance could be interpreted as remarkable cognitive performance on the part of the apes ceases to be impressive when the methodology or the conceptual background is scrutinized critically. This scrutiny consistently shows that their SWM and AWM subsystems show qualitative limitations in comparison to humans.

### 7.2.2. Comparison of group level performance in Homo versus Pan

Our rating of how experimental results relate chimpanzee performance to human performance is virtually the opposite of the results obtained through the hierarchical rating system of cognitive abilities devised by Carruthers (2013) that claims chimpanzees have essentially the same working memory capacity as do humans. Carruthers classifies working memory performance in terms of ascending levels of cognitive demand going from 1 to 8. We present his levels of cognitive demand in tabular form in Table 5. Carruthers only considers pure working memory and not effective working memory, and places increase in size of pure working memory at the beginning of his hierarchy. This leads him to assert, incorrectly as we discussed above in Section 6.7.3.2, that for chimpanzees PWMS\* = 4 = PWMS. In addition, he provides no reason for assuming that the working memory properties that are part of his

**Table 5**  
Posited Working Memory Hierarchy.

Level 1	No working memory
Level 2	PWMS = 2, but only in the absence of interference
Level 3	PWMS = 4, but only in the absence of interference
Level 4	Level 3 (not affected by interference); cannot be initiated <i>ab initio</i> through top-down initiation of WM
Level 5	Level 4 (can use top-down initiation); cannot use rehearsal to generate content for WM
Level 6	Level 5 (can use rehearsal); limited ability to transform WM representations into problem solving
Level 7	Level 6 (can transform representations); cannot use WM content outside current tasks
Level 8	Level 7 (use WM content outside current tasks); have WM comparable to that of humans

Level descriptions modified from Carruthers (2013).

Levels 4–8 only occur after PWMS = 4. Since PWMS develops in humans with maturation and PWMS = 4 for 3-year-old children (but see Cowan et al., 2011 for evidence that PWMS = 2 for 6–9-year-olds), Carruthers' assumption that Levels 4–8 require PWMS = 4 implies that none of the properties associated with these levels occurs in human development before PWMS = 4, which does not make sense. He has provided no reason for excluding the possibility that chimpanzees may have some of the properties of Levels 4–8 even with PWMS\* = 2, hence Carruthers' claim that great apes have Level 7 working memory, even if valid, does not justify his claim that PWMS = PWMS\*.

Carruthers assigns great apes to level 7 owing to their supposed computational working memory capacity for the requisite mental manipulation for insightful problem solving and planning that fully engages this capacity. This, of course, is linked to the holding and strategic sub-systems of working memory, as well as to the attentional sub-system that, at least in humans, enables a theory of mind, though this last aspect seems to be at best poorly developed in apes, as we have discussed above. Carruthers (2013) refers to some of the studies discussed in the previous section by primatologists and comparative psychologists who attribute remarkable prospective abilities to great apes, and he accepts *prima facie* the soundness of their interpretations, except for the accounts of number performance by Pan, even though our review of these claims has shown that the interpretations are not justified by the evidence. A more critical approach than that provided by Carruthers is needed to disentangle the strategic (SWM) and attentional (AWM) aspects of chimpanzee working memory and its computational power.

### 7.3. Elaboration of working memory without language in *Homo erectus* but not in Pan

We have identified substantive differences between WM in chimpanzees and in humans, including difference in PWMS versus PWMS\* and EWMS versus EWMS\*, leading to qualitative differences in the cognitive performance of humans in comparison to chimpanzees. This is in striking contrast, as we discuss above, to the review by Carruthers (2013) of working memory in non-human animals, non-human primates, and humans in which he concludes that WM in non-human primates is homologous with that of humans, that the size of PWMS is the same in humans, non-human primates and other mammals, and chimpanzees share with humans WM that includes all of the criteria through Level 7 in Table 5. If so, where does the difference in cognitive abilities between humans and chimpanzees come from? He answers this question by asserting that what is unique to humans is “our ability to vastly extend the topics and forms of reflective thinking in which we can engage by virtue of our capacity for mental rehearsal of speech” (Carruthers, 2013, p. 10377, *emphasis added*). This implies that prior to the development of language, hominin WM and chimpanzee WM were essentially the same. This is a testable proposition. Consider how the artifacts made by *Homo erectus*, dating from around 1.8 to 0.5 mya, hence before the origin of language, which probably occurred barely 200 kya (Knight, 2009), relate to WM. As one researcher puts it:

About 1.8 million years ago [before speech] ... *Homo erectus* strikes a rock against another while holding the idea of a shape in his mind. Through a series of skillful choices of angles and strikes, he imposes that shape to produce a hand axe ... it is unlikely – impossible, really – that another species could conceive and execute the complex planning inherent in [this scenario] ... There is something else at play: an ability to construct and carry out increasingly elaborate plans of action. (Coolidge et al., 2012, p.38).

Making a hand-axe involves a planning scenario that is striking in its complexity in comparison to a simple scenario for chimpanzees making termite sticks (Read and Andersson, 2019; Stout, 2011). Although comparative researchers have usually separated understanding relating to the physical world and the forces that govern it (physical cognition)

from the understanding of other minds (social cognition), there is an aspect of cognition feeding both – a common “machinery,” so to speak, that can be seen in the scenario for making a hand-axe. When a critical level of development (threshold) is reached through expansion of working memory, a qualitatively new way of processing information is enabled in a comparable way across all our complex cognitive abilities. When working memory can be activated by mentally entertaining the images/thoughts/feelings we ascribe to another person as well as one's own, and these representations can be kept distinct conceptually, it is possible to shift from one representation to the other to extract and/or compute information relevant to making appropriate social decisions. The simultaneous activation of alternative representations occurs when knapping a hand-axe. The artisan has in mind not only a representation of the final, intended product and its desired features, but may also have in mind an alternative artifact (e.g. a cutting implement) that he or she could possibly make, while simultaneously focusing on, and mentally representing, the stone substrate that he or she is working with but which does not yet correspond to either the hand-axe currently conceptualized by the artisan or the cutting tool that is also being conceived by the artisan. All these imagined representations and scenarios can be entertained by the artisan as he or she proceeds when working memory is powerful enough to entertain multiple spheres of reality or mental objects for which information is updated as the artisan proceeds, and each of which can be mentally accessed by the artisan at will. In this sense, theory of mind with its multiple representations and use of episodic memory, and prospective preparation of tools with its multiple representations, parallels the properties of working memory and this suggests that prospective and episodic memory probably appeared roughly at the same time and developed hand in hand.

The striking difference in complexity for the planning scenario for making a hand-axe and its dependence on an already elaborated WM in *Homo erectus* in comparison to that of a chimpanzee-like ancestor involves elaboration of WM that obviously does not derive from speech rehearsal, contrary to Carruthers' argument. Since these and other data (see Read and van der Leeuw, 2008) from hominin activity before the

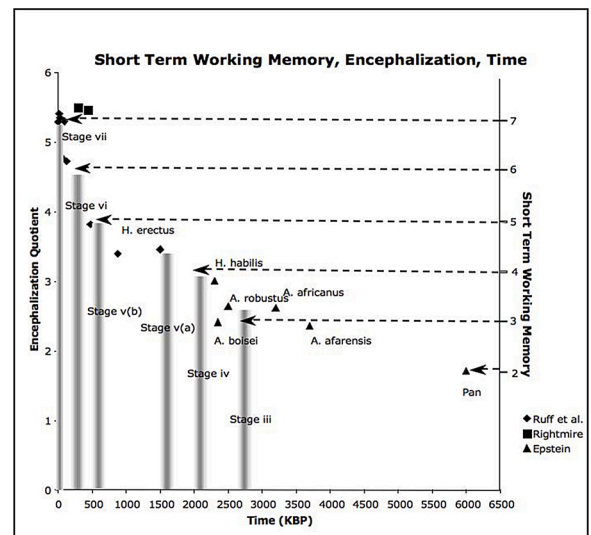


Fig. 12. Comparison of the size of working memory (right vertical axis) with change in the encephalization quotient (left vertical axis) through time in hominin phylogenetic groups and with increase in the conceptual dimensionality of tool production during hominid evolution divided into 7 stages (only Stages iii to vii relate to stone tool production). Vertical “fuzzy” bars identify the approximate time periods for the stages relating to stone tool production. Encephalization quotient based on EQ = brain mass/(11.22\*body mass<sup>0.76</sup>) (Martin, 1981). Stone tool data source: Read and van der Leeuw (2008). Fossil data sources: triangles, Epstein (2002, Rightmire (2004, Ruff et al. (1997). (Modified from Read and van der Leeuw, 2008, Fig. 2).

origin of syntactic forms of language all show that language is not the basis for more elaborated working memory in hominin ancestors to *Homo sapiens*, it is necessary to turn to what we understand about the evolutionary expansion and elaboration of WM in the hominins that was already playing out by the time of *Homo erectus* (see Fig. 12).

## 8. Relationship of expansion and elaboration of working memory to the evolution of *Homo sapiens*

The last common ancestor for *Homo* and *Pan* dates to around 6–8 mya, hence WM capacity would be, at most, the same as for *Pan* today, so assume WM capacity =  $2 \pm 1$  for the last common ancestor, meaning that during hominin evolution leading to *Homo sapiens*, EWM capacity increased from  $2 \pm 1$  to  $7 \pm 2$ . Assuming a linear increase in EWM capacity through time, evolutionary increase in EWM capacity corresponds to the hominin phylogeny leading to *Homo sapiens* as shown in Fig. 12. Striking is the correlation between stepwise increase in EWM capacity and major events in the hominin phylogeny leading to *Homo sapiens*.

We can also relate the increase in EWM capacity to archaeological evidence for qualitative changes marking different stages in the design and technological complexity of tool manufacture (see Read and van der Leeuw, 2008), beginning with tools used by chimpanzees and extending through the stone tools that are critical to hominin evolution leading to *Homo sapiens*. As shown in Table 6 and with the hominins, stone tool manufacturing evolves from the earliest technology for removal of flakes used as tools (Stage 2, 3.3 mya, found at Lomekwi, West Turkana, Kenya (Harmand et al., 2015)) through the complex blade technologies eventually appearing in most parts of the world (Stage 7, c. 50,000 BP). Columns 2–5 in Table 6 identify some of the observations linking stone tool manufacturing with EWM performance indexed by the size of EWM (Column 8) and the approximate date for the beginning of each stage (Column 9). Column 10 identifies newly appearing artifact categories corresponding to each stage. (For a more extensive discussion, see Read and van der Leeuw, 2008, pp. 1961–1964, 2015). As can be seen from Table 6 and Fig. 12, overall, there is an interrelated pattern incorporating cranial, technological, and behavioral changes associated with

increase in effective working memory capacity that is a central part of hominin evolution leading to *Homo sapiens*. But to detail that pattern would require another account.

## 9. Conclusions

Research over the past several decades has drastically expanded our understanding of the complexity and multi-facetedness of the behavior of the non-human primates, especially of the great apes, in general, and of the chimpanzees, in particular. This research challenges our understanding of what it means to be human and to what extent, and even if, there is a qualitative difference between ourselves and the non-human primate species biologically nearest to us. Working memory provides an avenue for exploring the degree and extent to which we share commonality with the non-human primates due its centrality and developmental trajectory from birth to adulthood regarding cognitive performance of individuals in our species. The breadth of this developmental trajectory enables a variety of ways that comparison may be made between the cognitive abilities of ourselves and the non-human primates. What these comparisons consistently show, regardless of the modality for the comparison (e.g., behaviors such as tool using, tool making, theory of mind, social behavior, and planning, or changes in brain architecture such as patterns of brain growth and association of cognitive abilities with brain regions), is a critical transition when going from WM =  $2 \pm 1$  to WM >  $2 \pm 1$ .

For the chimpanzees, WM =  $2 \pm 1$  is a stable equilibrium that suffices as the basis for a rich and wide-ranging array of behaviors. These behaviors were made possible by evolutionary changes in earlier non-human primates leading to the great apes. These changes began to downplay genetically endowed behavior in the evolution of the great apes through selection for individualistic behavior whose degree of occurrence and range of expression became comparable to what we find in our own species. Whether the comparison is tool making and using abilities, innovativeness in behavior, forming of social relations through extensive face-to-face interaction, rudiments of, at least, social traditions, ability to modulate behavior in accordance with pre-cursor mental capacities associated with the rubric, theory of mind, and the like, the

**Table 6**  
Evolution of Tool Complexity Compared to Change in Size of WM.

Stage	Concept	Action	Innovation	Dimensionality	Goal	Mode	WM size	Age BP	Example
1	Object attribute	Repetition possible	Functional attributes already present, can be enhanced	0	Use an object		1		Termite stick
1A	Relationship between objects	Objects used jointly	Using more than one object to fulfill task	0	Combine objects in a task		2		Nut cracking
2	Impose attribute	Repetition possible	Raw material modified to fulfill task	0	Detach flakes		2	3.3 mya	Lomekwi 3 <sup>a</sup>
3	Flaking	Repetition	Deliberate flaking, but without overall design	0: Incident angle < 90 °; conchoidal flaking	Shape flakes		3	2.6 mya	Lokalalei 2C <sup>b</sup>
4	Edge	Iteration: each flake controls the next	Débitage: flaking to create an edge on a core	1: Line of flakes creates partial boundary = edge	Shape core	1	4	2.0 mya	Oldowan chopper
5	Closed curve	Iteration: each flake controls the next	Débitage: flaking to create an edge determining a surface	2: Edges as generative elements of surfaces	Shape biface from edge	1	4.5	1.6 mya	Early hand axes
5A	Surface	Iteration: each flake controls the next	Façonnage: flaking of surface to make a shape	2: Surface intended elements, organized in relation to one another	Shape biface from surfaces	2	5	500 kya	Biface hand axes
6	Surface	Algorithm for flake removal	Control over location and angle of flaking to form surface	2: Surface of the flake brought under control but with shape constraint	Multiple production of tools	3	6	300 kya	Levallois technique
7	Intersection of planes	Recursive application of algorithm	Prismatic blade technology; repetitive production	3: Flake removal prepares core for next removal; no shape constraint on tool form	Serial production of tools	4	7	> 50 kya	Blade technologies

Modified from van der Leeuw (2020: Table 8.11).

<sup>a</sup> Harmand et al. (2015).

<sup>b</sup> Delagnes and Roche (2005).

cognitive capacities associated with  $WM = 2 \pm 1$  (see Figs. 5 and 6) are sufficient to provide a stable foundation for the occurrence and implementation of a wide range of cognitive abilities and associated behaviors in the chimpanzees. This is evidenced by relative stasis in the size of their working memory since the time of a last common ancestor with *Homo sapiens* to the present.

Attempts to argue that their WM is comparable to that of *Homo sapiens* are not supported by the evidence, as we have demonstrated, regardless of the modality for behavior being compared. At best, there may be an exceptional individual that displays unusual capacities in comparison to her/his compeers, just as there are exceptional humans that display unusual capacities in comparison to her/his compeers, but what we are concerned with is not the working memory capacity of the exceptional individual but the working memory capacities that characterize the species, whether it be *Pan troglodytes* or *Homo sapiens*. It is here that we find a qualitative difference – but not a rupturist change – in that once our hominin ancestors entered into an adaptive mode that included a selective advantage for cranial expansion via a feedback process between new modes of cultural adaptation and cranial volume (Markov and Markov, 2020), with the latter leading to expansion of both pure and effective working memory capacity, this enabled, over hundreds of thousands to millions of years, the development of a cultural and not just a biological mode of adaptation (Read, 2012b). Whether it be innovation in tool making technologies (see Fig. 12 and Table 6), innovation in social relations through the evolution of culturally constructed systems of kinship relations (Read, 2019), transformations in social relations from face-to-face to culturally constructed relation-based modes of social relations (Read, 2010, 2012b), expansion of group dynamics from being time and space bound to no longer being time and space bounded (Gamble, 2007, 2010), we find qualitative and not just quantitative differences between the biological adaptations of the non-human primates and the cultural adaptations of *Homo sapiens* (Lane et al., 2009; Read et al., 2009).

Too see how these changes were initiated through an increase in WM, consider first the increase in qualitatively different structural relations entailed by going from  $WM = 1 \pm 1$  to  $WM = 2 \pm 1$  and then to  $WM = 3 \pm 1$ . Each step is quantitatively small but the increase from  $WM = 2 \pm 1$  to  $WM = 3 \pm 1$  has a large qualitative consequence in comparison to the increase from  $WM = 1 \pm 1$  to  $WM = 2 \pm 1$ .

With  $WM = 1$ , hence instantiated by, for example, a single individual A, there is but one network formed by a single individual. With  $WM = 2$  and instantiated by, say, individuals A and B, there are four qualitatively different networks that can be formed from A and B: (1) A is isolated from B, (2) A is connected to B and B is connected to A, (3) A is connected to B but B is not connected to A, and (4) B is connected to A but A is not connected to B. With three persons A, B and C there are now  $4^3 = 64$  qualitatively possible different networks. In addition, the instantiation of  $WM = 1$  identifies the point of interest; the instantiation of  $WM = 2$  can identify a relation between A and B; and the instantiation of  $WM = 3$  can represent a relation from A to C computed from the relation of A to B and of B to C. The transformation from  $WM = 2$  to  $WM = 3$  is critical as it introduces the computational basis for the development of cultural ideational systems (Leaf and Read, 2012) that are central to the cultural adaptation of *Homo sapiens* in defining what we are as a species (see Read, 2012b).

We now outline a possible origin in early hominin evolution for the transformation going from  $WM = 2 \pm 1$  to  $WM = 3 \pm 1$  that initiated a feedback loop eventually leading to  $EWM = 7 \pm 2$  and affecting all behavioral aspects of hominin evolution leading to modern *Homo sapiens* (Read and van der Leeuw, 2008; see Fig. 12). Beginning around 6–8 mya with the hominin divergence from a common ancestor for *Pan* and *Homo*, to around 3 mya and phylogenetically marked by *Australopithecus* being replaced by *Homo*, a major evolutionary change was the development of an adaptation based on habitual bipedalism.

The habitual bipedalism of australopithecines plausibly favored manipulation of stone, including removal of stone flakes and possibly

their application to animal remains, between 3.5 and 3 mya (Harmand et al., 2015; McPherron et al., 2010). Initially, flakes could have been made by simply using one cobble to break up another (which only requires  $WM = 2$ ) and selecting a flake with a sharp edge from the debris. By 2.6–2.5 mya stone-flaking was frequent (Braun et al., 2019; Harris, 1986; Kimbel et al., 1996; Semaw et al., 1997, 2003), probably becoming habitual behavior (Shea, 2017), markedly different from stones battered by anthropoid primates (Arroyo et al., 2016; Toth et al., 2006), and perhaps attributable to an early evolutionary appearance in Africa of *Homo* which is claimed on the basis of a small number of fossil fragments from 2.8–2.3 mya (Kimbel et al., 1996; Villmoare et al., 2015). Between 2.6 and 2.3 mya African assemblages of flaked stones have permitted the refitting of excavated pieces that show repetitive striking where siliceous stones afforded surface angles  $\leq 90^\circ$  (Delagnes and Roche, 2005). That “affordance”, which confounds apes trained to flake stones, favors creation of overlapping, shallow, concave (“conchoidal”) scars, thereby producing cutting edges on “chopping tools” by consecutive removal of sharp flakes with (“bulbar”) convexities. With conchoidal flaking, “... toolmaking represents a true break from the capacities of nonhuman hominids ...” (Hovers, 2012, p. 59).

There is little doubt that stone artefacts were being used to extract edible foodstuff from bioenergetic resources available in animals and plants, and that conchoidal flaking facilitated such use. It is worth commenting here that between 4 and 2 mya, in relation to body-size there is an inverse relationship, from australopithecines to early Pleistocene *Homo*, between brain-size, which increases, and molar dimensions which decrease (Macho and Wood, 1995; McHenry, 2002; Pilbeam and Gould, 1974), which may reflect a growing consumption of animal tissues, though from 2–1.5 mya, when early *Homo* body-size was approaching that of modern people, brain-size increased faster than tooth-sizes decreased (Gómez-Robles et al., 2017), which is physiologically reasonable because the bioenergetic requirements of a large body and a large brain cannot be met by chewing with very small teeth in the absence of a sophisticated technology for preparing foods by concentrating them into small volumes that nevertheless can satisfy our voracious metabolic demands.

Conchoidal flaking on a regular basis by adult group members depends on  $WM = 3$  since, just as nut cracking involves three objects manipulated in a fixed sequential order and requires  $WM = 3$ , removing a flake through conchoidal flaking requires a core nodule, a hammerstone and a particular percussion angle, all to be controlled in a fixed sequential order, hence  $WM = 3$  is required. Though both require  $WM = 3$ , conchoidal flaking differs from nut cracking in the following important way regarding evolutionary consequences. The sharp cutting edges of flakes and chopping tools enabled fast and efficient extraction of fat-rich soft tissues, bone marrow, and meat from the carcasses of herbivorous mammals, which made possible enrichment of the diet of early *Homo* through dismemberment of carcasses, whether scavenged after dying from natural causes or after having been killed by other predators. This placed them in competition with fierce carnivores for exploiting animal foodstuffs, but efficient extraction of meat and soft tissues by early humans with stone tools, that could also be thrown as weapons at animals and vultures, and bring down accessible prey, afforded them a competitive edge over their rivals, both figuratively and literally. This gave evolving *Homo* increased adaptive fitness through exploitation of conchoidal flaking. In contrast, nuts suitable for nut cracking are still not part of the diet of all chimpanzees, from which it is reasonable to infer that nut cracking does not have the same relationship to evolving adaptive fitness as does conchoidal flaking of stone artefacts by our human forebearers, hence it is simply one of several ways that chimpanzees exploit food resources.

Evolution in working memory leading to modern *Homo sapiens* would have begun from a common ancestor with  $PWMS = 2 \pm 1 = EWMS$  and then evolved to  $PWMS = 4 \pm 1$  (Cowan, 2001) and  $EWMS = 7 \pm 2$  (Miller, 1956) that we find in *Homo sapiens* today. This involved both evolutionary change in the architecture of WM, as indicated by the



increase in PWMS from  $2 \pm 1$  to  $4 \pm 1$ , and in the implementation of that architecture, as indicated by the increase in EWMS from  $2 \pm 1$  to  $7 \pm 2$ . The greater increase in the size of effective WM in comparison to the size of pure WM indicates that implementation of the architecture became increasingly important in the evolution of working memory leading to modern *Homo sapiens*. This contrasts sharply with stasis in *Pan* for both PWMS and EWMS over the same time period and suggests that much of the difference between humans and our closest living relatives in relation to WM computational power results from increase in the size of effective WM.

We find that cognitive mechanisms regarded as “distinctively human” or only nascent in other species, such as

... (1) mechanisms that are specialized for dealing with the inanimate world, such as causal understanding; (2) faculties that are equally likely to process animate (social) and inanimate (asocial) events, such as episodic memory; and (3) various forms of cognition specialized for dealing with social stimuli, such as face processing, imitation, and mindreading ... (Heyes, 2018, p. 48),

are distinguished by our having cognitive abilities such as “causal understanding” and “mindreading” that depend on  $HWM = 3$  or more for effective implementation. It is not the quantitative size of WM derived directly from the architecture of the brain (PWMS) which matters most, but the qualitative changes enabled through the implementation of that architecture (EWMS). What appears to distinguish humans is their capacity for implementing strategies that give access to the computational power of their WM and this provided *Homo* with a computational power vastly exceeding what occurs in the non-human primates. We suggest, then, that what differentiates humans from other animals is our capacity to “harness the power,” as we metaphorically put it, derived from pure WM through a highly developed metacognitive ability that had the consequence of transforming PWMS =  $4 \pm 1$  into EWMS  $7 \pm 2$ . This also brings to mind the metaphor of the “inner eye,” proposed by Nicholas Humphrey (1986), that what characterizes human beings is the capacity to explore our mind through being aware of our own thought processes and thereby transforming individual metacognition into group level shared cognition.

In conclusion, the transition to  $WM = 3 \pm 1$  was a quantitative change that initiated a feedback loop that enabled causal relations and reasoning of the form “A in the context of B, gives rise to C,” hence the ability to recognize contingency in causal relations that otherwise are of the more limited form “A gives rise to B.” The qualitative changes enabled by quantitative changes leading to expanded working memory laid the foundation for the inclusion in our species of a cultural adaptation that introduced capabilities in our species that are unavailable through the biological adaptations of the non-human primates.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References

- Alexander, G.E., Goldman, P.S., 1978. Functional development of the dorsolateral prefrontal cortex: an analysis utilizing reversible cryogenic depression. *Brain Res.* 143, 233–249.
- Alloway, T.P., 2007. Working memory, reading, and mathematical skills in children with developmental coordination disorder. *J. Exp. Child Psychol.* 96, 20–36. <https://doi.org/10.1016/j.jecp.2006.07.002>.
- Alloway, T.P., Alloway, R.G., 2010. Investigating the predictive roles of working memory and IQ in academic attainment. *J. Exp. Child Psychol.* 106, 20–29. <https://doi.org/10.1016/j.jecp.2009.11.003>.
- Alloway, T.P., Archibald, L.M., 2008. Working memory and learning in children with developmental coordination disorder and specific language impairment. *J. Learn. Disabil.* 41, 251–262. <https://doi.org/10.1177/0022219408315815>.
- Alp, I.E., 1989. Mental Capacity and Working Memory in 1 to 3 Year Olds. Unpublished Doctoral Dissertation. York University, Toronto.
- Alp, I.E., 1994. Measuring the size of working memory in very young children: the imitation sorting task. *Int. J. Behav. Dev.* 17, 125–141. <https://doi.org/10.1177/016502549400700108>.
- Apperly, I.A., Butterfill, S.A., 2009. Do humans have two systems to track beliefs and belief-like states? *Psychol. Rev.* 116 (4), 953–970.
- Arroyo, A., Hirata, S., Matsuzawa, T., de la Torre, I., 2016. Nut cracking tools used by captive chimpanzees (*Pan troglodytes*) and their comparison with Early Stone Age percussive artefacts from Olduvai Gorge. *PLoS One* 11 (11), e166788. <https://doi.org/10.1371/journal.pone.0166788>.
- Ashkenazi, S., 2016. Enumeration processes under attack: the role of working memory in subitizing and serial counting. *J. Forensic Psychol.* 1 (1) <https://doi.org/10.4172/2475-319X.1000103>.
- Baddeley, A.D., 1986. *Working Memory*. Clarendon Press, Oxford.
- Baddeley, A.D., 1996. Exploring the central executive. *Q. J. Exp. Psychol.* 49A (1), 5–28.
- Baddeley, A.D., 2000. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences* 4, 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2).
- Baddeley, A.D., 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4 (10), 829–839.
- Baddeley, A.D., 2012. Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>.
- Balter, M., 2010. Did working memory spark creative culture? *Science* 328, 160–163. <https://doi.org/10.1126/science.328.5975.160>.
- Baron-Cohen, S., Leslie, A.M., Frith, U., 1985. Does the autistic child have a “theory of mind”? *Cognition* 21, 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8).
- Bartz, A.E., 1962. Eye-movement latency, duration, and response time as a function of angular displacement. *J. Exp. Psychol.* 64, 318–324. <https://doi.org/10.1037/h0043475>.
- Bayles, K.A., 2003. Effects of working memory deficits on the communicative functioning of Alzheimer’s dementia patients. *J. Commun. Disord.* 36 (3), 209–219. [https://doi.org/10.1016/S0021-9924\(03\)0020-0](https://doi.org/10.1016/S0021-9924(03)0020-0).
- Beaman, C.P., 2010. Working memory and working attention. What could possibly evolve? *Curr. Anthropol.* 51 (Suppl. 1), S21–S38. <https://doi.org/10.1086/650297>.
- Beran, M.J., Beran, M.M., 2004. Chimpanzees remember the results of one-by-one addition of food items to sets over extended time periods. *Psychol. Sci.* 15 (2), 94–99. <https://doi.org/10.1111/j.0963-7214.2004.01502004.x>.
- Bering, J.M., Bjorklund, D.F., Ragan, P., 2000. Deferred imitation of object-related actions in human-reared juvenile chimpanzees and orangutans. *Dev. Psychobiol.* 36, 218–232. [https://doi.org/10.1002/\(SICI\)1098-2302\(200004\)36:3%3C218::AID-DEV5%3E3.0.CO;2-K](https://doi.org/10.1002/(SICI)1098-2302(200004)36:3%3C218::AID-DEV5%3E3.0.CO;2-K).
- Boesch, C., 1994. Cooperative hunting in wild chimpanzees. *Anim. Behav.* 48, 653–667. <https://doi.org/10.1006/anbe.1994.1285>.
- Boesch, C., 2002. Cooperative hunting roles among Tai chimpanzees. *Hum. Nat.* 13, 27–46. <https://doi.org/10.1007/s12110-002-1013-6>.
- Boesch, C., Boesch, H., 1983. Optimization of nut-cracking with natural hammers by wild chimpanzees. *Behaviour* 83, 265–286. <https://doi.org/10.1007/BF02382388>.
- Boesch, C., Head, J., Robbins, M., 2009. Complex tool sets for honey extraction among chimpanzees in Loango National Park, Gabon. *J. Hum. Evol.* 56, 560–569. <https://doi.org/10.1016/j.jhevol.2009.04.001>.
- Botvinick, M.M., Wang, J., Gowan, E., Roy, S., Bastianen, C., Mayo, J.P., Houk, C., 2009. An analysis of immediate serial recall performance in a macaque. *Anim. Cogn.* 12, 671–678. <https://doi.org/10.1007/s10071-009-0226-z>.
- Brady, R.J., Hampton, R.R., 2018. Nonverbal working memory for novel images in Rhesus monkeys. *Curr. Biol.* 28 (24), 3903–3910. <https://doi.org/10.1016/j.cub.2018.10.025> e3.
- Bräuer, J., Call, J., 2015. Apes produce tools for future use. *Am. J. Primatol.* 77, 254–263. <https://doi.org/10.1002/ajp.12341>.
- Braun, D.R., Aldeias, V., Archer, W., Arrowsmith, J.R., Baraki, N., Campisano, C.J., Deino, A.L., DiMaggio, E.N., Dupont-Nivet, G., Engda, B., Feary, D.A., Garello, D.I., Kerfelew, Z., McPherron, S.P., Patterson, D.B., Reeves, J.S., Thompson, J.C., Reed, K.E., 2019. Earliest known Oldowan artifacts > 2.58 Ma from Ledi-Geraru, Ethiopia, highlight early technological diversity. *Proc. Natl. Acad. Sci. U. S. A.* 116 (24), 11712–11717. <https://doi.org/10.1073/pnas.1820177116>.
- Brewer, S.M., McGrew, W.C., 1990. Chimpanzee use of a tool-set to get honey. *Folia Primatol.* 54, 100–104. <https://doi.org/10.1159/000156429>.
- Bull, R., Scerif, G., 2012. Executive functioning as a predictor of children’s mathematics ability: shifting, inhibition, and working memory. *Dev. Neuropsychol.* 19, 273–293. [https://doi.org/10.1207/S15326942DN1903\\_3](https://doi.org/10.1207/S15326942DN1903_3).
- Burgoyne, A.P., Engle, R.W., 2020. Attentional control: a cornerstone of higher-order cognition. *Curr. Dir. Psychol. Sci.* 29 (6), 624–630. <https://doi.org/10.1177/09637214200969371>.
- Buxhoevedn, D., Casanova, M.F., 2002. The minicolumn and evolution of the brain. *Brain Behav. Evol.* 60, 125–151. <https://doi.org/10.1159/000065935>.
- Call, J., Tomasello, M., 2008. Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* 12, 187–192. <https://doi.org/10.1016/j.tics.2008.02.010>.
- Carlson, S.M., Moses, L.J., Breton, C., 2002. How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant Child Dev.* 11, 73–92. <https://doi.org/10.1002/icd.298>.
- Carpenter, M., Call, J., 2013. How joint is the joint attention of apes and human infants? In: Metcalfe, J., Terrace, H.S. (Eds.), *Agency and Joint Attention*. Oxford University Press, New York, pp. 49–61.
- Carpenter, M., Nagell, K., Tomasello, M., 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, 1–143. <https://doi.org/10.2307/1166214>.
- Carruthers, P., 2013. Evolution of working memory. *Proc. Natl. Acad. Sci. U. S. A.* 110, 10371–10378. <https://doi.org/10.1073/pnas.1301195110>.

- Casey, B.J., Tottenham, N., Liston, C., Durston, S., 2005. Imaging the developing brain: what have we learned about cognitive development? *Trends Cogn. Sci.* 9 (3), 104–110. <https://doi.org/10.1016/j.tics.2005.01.011>.
- Castro-Caldas, A., Petersson, K.M., Reis, A., Stone-Elander, S., Ingvar, M., 1998. The illiterate brain. Learning to read and write during childhood influences the functional organization of the human brain. *Brain* 121, 1053–1063.
- Chai, W.J., Hamid, A.I.A., Abdullah, J.M., 2018. Working memory from the psychological and neurosciences perspectives: a review. *Front. Psychol.* 9, 401. <https://doi.org/10.3389/fpsyg.2018.00401>.
- Christophel, T.B., Klink, P.C., Spitzer, B., Roelfsema, P.R., Haynes, J.D., 2017. The distributed nature of working memory. *Trends Cogn. Sci.* 21, 111–124. <https://doi.org/10.1016/j.tics.2016.12.007>.
- Conklin, H.M., Luciana, M., Hooper, C.J., Yarger, R.S., 2007. Working memory performance in typically developing children and adolescents: behavioral evidence of protracted frontal lobe development. *Dev. Neuropsychol.* 31 (1), 103–128. [https://doi.org/10.1207/s15326942dn3101\\_6](https://doi.org/10.1207/s15326942dn3101_6).
- Conway, A.R.A., Kane, M., Engle, R.W., 2003. Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences* 7, 547–552. <https://doi.org/10.1016/j.tics.2003.10.005>.
- Cook, P., Wilson, M., 2010. Do young chimpanzees have extraordinary working memory? *Psychon. Bull. Rev.* 17, 599–600. <https://doi.org/10.3758/PBR.17.4.599>.
- Coolidge, F.L., Wynn, T., 2001. Executive functions of the frontal lobes and the evolutionary ascendancy of Homo sapiens. *Cambridge Archaeol. J.* 11, 255–260. <https://doi.org/10.1017/S0959774301000142>.
- Coolidge, F.L., Wynn, T., Overmann, R.S., 2012. The evolution of working memory. In: Alloway, T.J.P., Alloway, F.G. (Eds.), *Working Memory: The Connected Intelligence*. Psychology Press, New York, pp. 37–40.
- Corballis, M.C., 2011. *The Recursive Mind. The Origins of Human Language, Thought, and Civilization*. Princeton University Press, Princeton.
- Corrêa, L.M.S., 1995. An alternative assessment of children's comprehension of relative clauses. *J. Psycholinguist. Res.* 24, 183–203. <https://doi.org/10.1007/BF02145355>.
- Cowan, N., 1988. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol. Bull.* 104, 163–191. <https://doi.org/10.1037/0033-2909.104.2.163>.
- Cowan, N., 1995. *Attention and memory: an integrated framework*. Oxford Psychology Series 26. Oxford University Press, Oxford and New York.
- Cowan, N., 1999. An embedded-process model of working memory. In: Miyake, A., Shah, P. (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge University Press, Cambridge, pp. 62–101.
- Cowan, N., 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185. <https://doi.org/10.1017/S0140525X01003922>.
- Cowan, N., 2008. What are the differences between long-term, short-term, and working memory? *Prog. Brain Res.* 169, 323–338. [https://doi.org/10.1016/S0079-6123\(07\)00020-9](https://doi.org/10.1016/S0079-6123(07)00020-9).
- Cowan, N., 2010. The magical mystery four: how is working memory capacity limited, and why? *Curr. Dir. Psychol. Sci.* 19 (1), 51–57. <https://doi.org/10.1177/0963721409359277>.
- Cowan, N., Alloway, T.P., 2008. The development of working memory. In: Cowan, N. (Ed.), *Development of Memory in Childhood*, 2nd ed. Psychology Press, Hove, pp. 303–342.
- Cowan, N., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S.M., Hismjatullina, A., Conway, A.R.A., 2005. On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51 (1), 42–100. <https://doi.org/10.1016/j.cogpsych.2004.12.001>.
- Cowan, N., Fristoe, N.M., Elliott, E.M., Brunner, R.P., Saults, J.S., 2006. Scope of attention, control of attention, and intelligence in children and adults. *Mem. Cognit.* 34, 1754–1768. <https://doi.org/10.3758/BF03195936>.
- Cowan, N., Morey, C.C., Chen, Z., 2007. The legend of the magical number seven. In: Della Sala, S. (Ed.), *Tall Tales about the Mind and Brain: Separating Fact from Fiction*. Oxford University Press, Oxford, pp. 45–59.
- Cutini, S., Bonato, M., 2012. Subitizing and visual short-term memory in human and non-human species: a common shared system? *Front. Psychol.* 3, 469–473. <https://doi.org/10.3389/fpsyg.2012.00469>.
- D'Esposito, M., Postle, B.R., 2015. The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* 66, 115–142. <https://doi.org/10.1146/annurev-psych-010814-015031>.
- Darki, F., Klingberg, T., 2015. The role of fronto-parietal and fronto-striatal networks in the development of working memory: a longitudinal study. *Cereb. Cortex* 25 (6), 1587–1595. <https://doi.org/10.1093/cercor/bht352>.
- Dekaban, A.S., 1978. Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Ann. Neurol.* 4, 345–356. <https://doi.org/10.1002/ana.410040410>.
- Delagnes, A., Roche, H., 2005. Late Pliocene hominid knapping skills: the case of Lokalelei 2C, West Turkana, Kenya. *J. Hum. Evol.* 48, 435–472. <https://doi.org/10.1016/j.jhevol.2004.12.005>.
- DeSilva, J., Lesnik, J., 2006. Chimpanzee neonatal brain size: implications for brain growth in *Homo erectus*. *J. Hum. Evol.* 51, 207e212. <https://doi.org/10.1016/j.jhevol.2006.05.006>.
- Diamond, A., 1990. Developmental time course in human infants and infant monkeys, and the neural bases of, inhibitory control in reaching. *Ann. N. Y. Acad. Sci.* 608, 637–669. <https://doi.org/10.1111/j.1749-6632.1990.tb48913.x> discussion 669–676.
- Diamond, A., 2013. Executive functions. *Annu. Rev. Psychol.* 64, 135–168.
- Diamond, A., Doar, B., 1989. The performance of human infants on a measure of frontal cortex function, the delayed-response task. *Dev. Psychobiol.* 22, 271–294. <https://doi.org/10.1002/dev.420220307>.
- Diamond, A., Prevor, M.B., Callender, G., Druin, D.P., 1997. Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monogr. Soc. Res. Child Dev.* 62, 206. <https://doi.org/10.2307/1166208>.
- Dumontheil, I., Klingberg, T., 2012. Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cereb. Cortex* 22, 1078–1085. <https://doi.org/10.1093/cercor/bhr175>.
- Elmore, L.C., Ma, W.J., Magnotti, J.F., Leising, K.J., Passaro, A.D., Katz, J.S., Wright, A.A., 2011. Visual short term memory compared in Rhesus monkeys and humans. *Curr. Biol.* 21, 975–979. <https://doi.org/10.1016/j.cub.2011.04.031>.
- Engelmann, J.M., Herrmann, E., Tomasello, M., 2012. Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS One* 7, e48433. <https://doi.org/10.1371/journal.pone.0048433>.
- Engle, R.W., 2002. Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. <https://doi.org/10.1111/1467-8721.00160>.
- Engle, R.W., Tuholski, S.W., Laughlin, J., Conway, A.R.A., 1999. Working memory, short-term memory and general fluid intelligence: a latent variable model approach. *J. Exp. Psychol. Gen.* 128, 309–331. <https://doi.org/10.1037/0096-3445.128.3.309>.
- Epstein, H.T., 2002. Evolution of the reasoning brain. *Behav. Brain Sci.* 25, 408–409. <https://doi.org/10.1017/S0140525X02200777>.
- Feigenson, L., 2007. Continuity of format and computation in short-term memory development. In: Oakes, M., Bauer, P.J. (Eds.), *Short- and Long-Term Memory in Infancy: Taking the First Steps Towards Memory*. Oxford University Press, New York, pp. 51–74.
- Feigenson, L., Carey, S., 2005. On the limits of infants' quantification of small object arrays. *Cognition* 97, 295–313. <https://doi.org/10.1016/j.cognition.2004.09.010>.
- Fitri, F., Fithrie, A., Rambe, A.S., 2020. Association between working memory impairment and activities of daily living in post-stroke patients. *Med Glas (Zenica)* 17 (2), 433–438.
- Funahashi, S., 2017. Working memory in the prefrontal cortex. *Brain Sci.* 7 (5), 49. <https://doi.org/10.3390/brainsci7050049>.
- Fuster, J.M., 2001. The prefrontal cortex – an update: time is of the essence. *Neuron* 30, 319–333. [https://doi.org/10.1016/S0896-6273\(01\)00285-9](https://doi.org/10.1016/S0896-6273(01)00285-9).
- Fuster, J.M., 2002. Frontal lobe and cognitive development. *J. Neurocytol.* 31, 373–385. <https://doi.org/10.1023/a:1024190429920>.
- Fuster, J.M., 2015. *The Prefrontal cortex*, 5th ed. Elsevier Academic Press, Amsterdam.
- Gamble, C., 2007. *Origins and Revolutions: Human Identity in Earliest Prehistory*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511618598>.
- Gamble, C., 2010. Technologies of separation and the evolution of social extension. In: Dunbar, R.I.M., Gamble, C., Gowlett, J. (Eds.), *Social Brain, Distributed Mind*. Oxford University Press, Oxford, pp. 17–42. <https://doi.org/10.5871/bacad/9780197264522.003.0002>.
- Gathercole, S.E., 2008. Working memory in the classroom. *Psychologist* 21, 382–385.
- Gathercole, S.E., Pickering, S.J., Ambridge, B., Wearing, H., 2004. The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* 40, 177–190. <https://doi.org/10.1037/0012-1649.40.2.177>.
- Gathercole, S.E., Alloway, T.P., Willis, C.S., Adams, A.M., 2006. Working memory in children with reading disabilities. *J. Exp. Child Psychol.* 93, 265–281. <https://doi.org/10.1016/j.jecp.2005.08.003>.
- Geary, D.C., Hoard, M.K., Hamson, C.O., 1999. Numerical and arithmetical cognition: patterns of functions and deficits in children at risk for a mathematical disability. *J. Exp. Child Psychol.* 74, 213–239. <https://doi.org/10.1006/jecp.1999.2515>.
- Gergely, G., Bekkering, H., Kiral, I., 2002. Developmental psychology: rational imitation in preverbal infants. *Nature* 415, 755. <https://doi.org/10.1038/415755a>.
- Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent III, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U. S. A.* 101, 8174–8179. <https://doi.org/10.1073/pnas.0402680101>.
- Goldberg, E., 2002. *The Executive Brain: Frontal Lobes and the Civilized Mind*. Oxford University Press, London.
- Goldman, P.S., Alexander, G.E., 1977. Maturation of prefrontal cortex in the monkey revealed by local reversible cryogenic depression. *Nature* 267 (5612), 613–615. <https://doi.org/10.1038/267613a0>.
- Gomez-Perez, E., Ostrosky-Solis, F., 2006. Attention and memory evaluation across the life span: heterogeneous effects of age and education. *J. Clin. Exp. Neuropsychol.* 28, 477–494.
- Gómez-Robles, A., Smaers, J.B., Holloway, R.L., Polly, P.D., Wood, B.A., 2017. Brain enlargement and dental reduction were not linked in hominin evolution. *Proc. Natl. Acad. Sci. U. S. A.* 114 (3), 468–473. <https://doi.org/10.1073/pnas.1608798114>.
- Goodall, J., 1998. Learning from the chimpanzees: a message humans can understand. *Science* 282, 2184–2185. <https://doi.org/10.1126/science.282.5397.2184>.
- Gray, C.R., Gummerman, K., 1975. The enigmatic eidetic image: a critical examination of methods, data, and theories. *Psychol. Bull.* 82, 383–407. <https://doi.org/10.1037/0033-2909.82.3.383>.
- Hannah, A.C., McGrew, W.C., 1987. Chimpanzees using stones to crack open oil palm nuts in Liberia. *Primates* 28 (1), 31–46. <https://doi.org/10.1007/BF02382181>.
- Hanus, D., Call, J., 2007. Discrete quantity judgments in the great ape (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus*): The effect of presenting whole sets versus item-by-item. *J. Comp. Psychol.* 121, 241–249. <https://doi.org/10.1037/0735-7036.121.3.241>.
- Hanus, D., Mendes, N., Tennie, C., Call, J., 2011. Comparing the performances of apes (*Gorilla*, *Pan troglodytes*, *Pongo pygmaeus*) and human children (*Homo sapiens*) in the

- floating peanut task. *Public Lib. Sci. One PLoS ONE* 6 (6), e19555. <https://doi.org/10.1371/journal.pone.0019555>.
- Hare, B., 2001. Can competitive paradigms increase the validity of experiments on primate social cognition? *Anim. Cogn.* 4, 269–280. <https://doi.org/10.1007/s100710100084>.
- Hare, B., Call, J., Agnetta, B., Tomasello, M., 2000. Chimpanzees know what conspecifics do and do not see. *Anim. Behav.* 59, 771–785. <https://doi.org/10.1006/anbe.1999.1377>.
- Hare, B., Call, J., Tomasello, M., 2001. Do chimpanzees know what conspecifics know? *Anim. Behav.* 61, 139–151. <https://doi.org/10.1006/anbe.2000.1518>.
- Harmand, S., Lewis, J.E., Feibel, C.S., Lepre, C.J., Prat, S., Lenoble, A., Boës, X., Quinn, R. L., Brenet, M., Arroyo, A., Taylor, N., Clément, S., Daver, G., Brugal, J.-P., Leakey, L., Mortlock, R.A., Wright, J.D., Lokorodi, S., Kirwa, C., Kent, D.V., Roche, H., 2015. 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature* 521, 310–315. <https://doi.org/10.1038/nature14464>.
- Harris, J.W.K., 1986. Découverte de matériel archéologique oldowayen dans le Rif de l'Afar. *L'anthropologie* 90, 339–357.
- Hauser, M.D., Carey, S., Hauser, L.B., 2000. Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proc. R. Soc. Series B* 267, 829–833. <https://doi.org/10.1098/rspb.2000.1078>.
- Hayashi, M., Matsuzawa, T., 2003. Cognitive development in object manipulation by infant chimpanzees. *Anim. Cogn.* 6, 225–233. <https://doi.org/10.1007/s10071-003-0185-8>.
- Henley, T., Povinelli, D.J., 2020. Seeing through <seeing through>: an analysis of Kano et al. (2019). *Anim. Behav. Cogn.* 7 (4), 658–660. <https://doi.org/10.26451/abc.07.04.11.2020>.
- Herrmann, E., Call, J., Hernández-Lloreda, M.V., Hare, B., Tomasello, M., 2007. Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317, 1360–1366.
- Herrmann, E., Hare, B., Call, J., Tomasello, M., 2010. Differences in the cognitive skills of bonobos and chimpanzees. *PLoS One* 5 (8), e12438. <https://doi.org/10.1371/journal.pone.0012438>.
- Herrmann, E., Keupp, S., Hare, B., Vaish, A., Tomasello, M., 2013. Direct and indirect reputation formation in nonhuman great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla*, *Pongo pygmaeus*) and human children (*Homo sapiens*). *J. Comp. Psychol.* 127 (1), 63–75. <https://doi.org/10.1037/a0028929>.
- Heyes, C.M., 1998. Theory of mind in nonhuman primates. *Behav. Brain Sci.* 21, 101–148. <https://doi.org/10.1017/s0140525x98000703>.
- Heyes, C.M., 2018. *Cognitive Gadgets: the Cultural Evolution of Thinking*. Harvard University Press. <https://doi.org/10.4159/9780674985155>.
- Horschler, D.J., MacLean, E.L., Santos, L.R., 2020. Do non-human primates really represent others' beliefs? *Trends Cogn. Sci.* 24 (8), 594–605.
- Hovers, E., 2012. Invention, reinvention and innovation: the makings of Oldowan lithic technology. In: Elias, S. (Ed.), *Origins of Human Innovation and Creativity*. Elsevier, Amsterdam, pp. 51–68.
- Hughes, C., Devine, R.T., Ensor, R., Koyasu, M., Mizokawa, A., Lecce, S., 2014. Lost in translation? Comparing British, Japanese, and Italian children's theory-of-mind performance. *Child Dev. Res.* 2014. <https://doi.org/10.1155/2014/893>.
- Humphrey, N., 1986. *The Inner Eye*. Faber & Faber, London.
- Humphrey, N., 2012. This chimp will kick your ass at memory games – but how the hell does he do it? *Trends Cogn. Sci.* 16, 353–355. <https://doi.org/10.1016/j.tics.2012.05.002>.
- Hunt, K.D., 2020. The mind of the chimpanzee. In: Lonsdorf, E.V., Ross, S.R., Matsuzawa, T. (Eds.), *Chimpanzee: Lessons from Our Sister Species*. Cambridge University Press, Cambridge, pp. 323–356.
- Inoue, S., Matsuzawa, T., 2007. Working memory of numerals in chimpanzees. *Curr. Biol.* 17, R1004–R1005. <https://doi.org/10.1016/j.cub.2007.10.027>.
- Inoue-Nakamura, N., Matsuzawa, T., 1997. Development of stone tool use by wild chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 3 (2), 159–173. <https://doi.org/10.1037/0735-7036.111.2.159>.
- Jaensch, E.R., 1930. Eidetic imagery and typological methods of investigation: their importance for the psychology of childhood, the theory of education, and general psychology. *International Library of Psychology* 42, *Cognitive Psychology* 5. Kegan Paul, London.
- Johnson, M.H., 2001. Functional brain development in humans. *Nat. Rev. Neurosci.* 2, 475–483. <https://doi.org/10.1159/000276477>.
- Johnson, J., Fabian, V., Pascual-Leone, J., 1989. Quantitative hardware stages that constrain language development. *Hum. Dev.* 32, 245–271.
- Káldy, Z., Leslie, A.M., 2003. Identification of objects in 9-month-old infants: integrating “what” and “where” information. *Dev. Sci.* 6, 360–373. <https://doi.org/10.1111/1467-7687.00290>.
- Kaminski, J., Call, J., Tomasello, M., 2008. Chimpanzees know what others know, but not what they believe. *Cognition* 109, 224–234. <https://doi.org/10.1016/j.cognition.2008.08.010>.
- Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., Call, J., 2019. Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proc. Natl. Acad. Sci. U. S. A.* 116 (42), 20904–20909. <https://doi.org/10.1073/pnas.1910095116>.
- Karg, K., Schmelz, M., Call, J., Tomasello, M., 2015. The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Anim. Behav.* 105, 211–221. <https://doi.org/10.1016/j.anbehav.2015.04.028>.
- Karg, K., Schmelz, M., Call, J., Tomasello, M., 2016. Differing views: Can chimpanzees do Level 2 perspective-taking? *Anim. Cogn.* 19, 555–564. <https://doi.org/10.1007/s10071-016-0956-7>.
- Kawai, N., Matsuzawa, T., 2000. Numerical memory span in a chimpanzee. *Nature* 403, 39–40. <https://doi.org/10.1038/47405>.
- Ken-Ichiro, T., Oyama, K., Nakamura, S., Iijima, T., 2016. Comparative overview of visuospatial working memory in monkeys and rats. *Front. Syst. Neurosci.* 10, 99. <https://www.frontiersin.org/article/10.3389/fnsys.2016.00099>.
- Kidd, E., Bavin, E.L., 2002. English-speaking children's comprehension of relative clauses: evidence for general-cognitive and language-specific constraints on development. *J. Psycholinguist. Res.* 31, 599–617. <https://doi.org/10.1023/A:1021265021141>.
- Kim, S., Sodian, B., Paulus, B., Senju, A., Okuno, A., Ueno, M., Itakura, S., Proust, J., 2020. Metacognition and mindreading in young children: a cross-cultural study. *Conscious. Cogn.* 85. <https://doi.org/10.1016/j.concog.2020.103017>.
- Kimbel, W.H., Walter, R.C., Johanson, D.C., Reed, K.E., Aronson, J.L., Assefa, Z., Marean, C.W., Eck, G.G., Bobe, R., Hovers, E., Rak, Y., Vondra, C., Yemana, T., York, D., Chen, Y., Eversen, N.M., Smith, P.E., 1996. Late Pliocene *Homo* and Oldowan tools from the Hadar formation (Kada Hadar Member), Ethiopia. *J. Hum. Evol.* 31 (6), 549–561. <https://doi.org/10.1006/jhev.1996.0079>.
- King, F.A., Yarbrough, C.J., Anderson, D.C., Gordon, P., Gould, K.G., 1988. *Primates. Science* 240, 1475–1482.
- Kliegel, M., Martin, M., McDaniel, M.A., Einstein, G.O., 2002. Complex prospective memory and executive control of working memory: a process mode. *Psychologische Beiträge* 44 (2), 303–319.
- Klingberg, T., Forssberg, S.J., Westerberg, H., 2002. Increased brain activity in frontal and parietal cortex underlies the development of Visuospace working memory capacity during childhood. *J. Cogn. Neurosci.* 14 (1), 1–10. <https://doi.org/10.1162/08992902317205276>.
- Knight, C., 2009. Introduction: perspectives on the evolution of language in Africa. In: Botha, R., Knight, C. (Eds.), *The Cradle of Language*. Oxford University Press, Oxford, pp. 1–15.
- Kosmidis, M.H., Tsapkinis, K., Folia, V., 2006. Lexical processing in illiteracy. *Cortex* 42, 1021–1027.
- Krachun, C., Carpenter, M., Call, J., Tomasello, M., 2009. A competitive nonverbal false belief task for children and apes. *Dev. Sci.* 12, 521–535. <https://doi.org/10.1111/j.1467-7687.2008.00793.x>.
- Krachun, C., Carpenter, M., Call, J., Tomasello, M., 2010. A new change-of contents false belief test: children and chimpanzees compared. *Int. J. Comp. Psychol.* 23, 145–165. Retrieved from. <https://escholarship.org/uc/item/68c0p8dk>.
- Krupenye, C., Kano, F., Hirata, S., Call, J., Tomasello, M., 2016. Great apes anticipate that other individuals will act according to false beliefs. *Science* 354, 110–114. <https://doi.org/10.1126/science.aaf8110>.
- Kulke, L., Rakoczy, H., 2017. How reliable and valid are anticipatory looking measures in theory of mind task? In: Rakoczy, H., Kulke (Chairs), L. (Eds.), *Are Implicit Theory of Mind Findings Robust? Some Doubts from Converging Non-Replications Across the Lifespan*. Symposium Conducted at the Society for Research in Child Development Biennial Meeting, Austin, Texas, April 2017.
- Lane, D., Maxfield, R.M., Read, D.W., van der Leeuw, S., 2009. From population to organization thinking. In: Lane, D., Pumain, D., van der Leeuw, S., West, G. (Eds.), *Complexity Perspectives on Innovation and Social Change*. Springer Verlag, Berlin, pp. 43–84. [https://doi.org/10.1007/978-1-4020-9663-1\\_2](https://doi.org/10.1007/978-1-4020-9663-1_2).
- Langer, J., 1986. *The Origins of Logic: One to Two Years*. Academic Press, New York.
- Langer, J., 2005. The primate phylogeny of cognitive ontogeny. In: Valsiner, J. (Ed.), *Heinz Werner and Developmental Science*. Springer, Berlin, pp. 307–319. [https://doi.org/10.1007/0-306-48677-6\\_19](https://doi.org/10.1007/0-306-48677-6_19).
- Langer, J., 2006. The heterochronic evolution of primate cognitive development. *Biol. Theory* 1 (1), 41–43. <https://doi.org/10.1162/biot.2006.1.1.41>.
- Lara, A.H., Wallis, J.D., 2015. The role of prefrontal cortex in working memory: a mini review. *Front. Syst. Neurosci.* 9, 173. <https://doi.org/10.3389/fnsys.2015.00173>.
- Leaf, M., Read, D., 2012. The organizational rubicon of the Upper Paleolithic. Chapter 3 in *Human Thought and Social Organization: Anthropology on a New Plane*. Lexington Books, Lanham, pp. 51–92. <https://doi.org/10.3389/fnsys.2015.00173>.
- Lewis, A., Berntsen, D., Call, J., 2019. Long term memory of past events in great apes. *Curr. Directions Psychol. Sci.* 28 (2), 117–123. <https://doi.org/10.1177/0963721418812781>.
- Lewis-Peacock, J.A., Cohen, J.D., Norman, K.A., 2016. Neural evidence of the strategic choice between working memory and episodic memory in prospective remembering. *Neuropsychologia* 93 (Pt A), 280–288. <https://doi.org/10.1016/j.neuropsychologia.2016.11.006>.
- Lind, J., Enquist, M., Ghirlanda, S., 2015. Animal memory: a review of delayed matching-to-sample data. *Behav. Processes* 117, 52–58.
- Luciana, M., Nelson, C.A., 1998. The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia* 36, 273–293. <https://doi.org/10.1016/j.jbeproc.2014.11.019>.
- Luciano, M., Wright, M.J., Smith, G.A., Geffen, G.M., Geffen, L.B., Martin, N.G., 2001. Genetic covariance among measures of information processing speed, working memory, and IQ. *Behav. Genet.* 31 (6), 581–592. <https://doi.org/10.1023/A:1013397428612>.
- Ma, W.J., Husain, M., Bays, P.M., 2014. Changing concepts of working memory. *Nat. Neurosci.* 17, 347–356.
- Macho, G.A., Wood, B.A., 1995. The role of timing in hominid dental evolution. *Evol. Anthropol.* 4 (1), 19–31. <https://doi.org/10.1002/evan.1360040105>.
- Majerus, S., Cowan, N., Péters, F., Van Calster, L., Phillips, C., Schrouff, J., 2016. Cross-modal decoding of neural patterns associated with working memory: evidence for attention-based accounts of working memory. *Cereb. Cortex* 26, 166–179.
- Majerus, S., Péters, F., Bouffier, M., Cowan, N., Phillips, C., 2017. The dorsal attention network reflects both encoding load and top-down control during working memory. *J. Cogn. Neurosci.* 30, 144–159. <https://doi.org/10.1093/cercor/bhu189>.

- Manrique, H.M., Walker, M.J., 2017. Early Evolution of Human Memory: Great Apes, Toolmaking, and Cognition. Palgrave Macmillan, Cham. <https://link.springer.com/book/10.1007/978-3-319-64447-9>.
- Markov, A.V., Markov, M.A., 2020. Runaway brain-culture coevolution as a reason for larger brains: exploring the “cultural drive” hypothesis by computer modeling. *Ecol. Evol.* 00, 1–19. <https://doi.org/10.1002/ece3.6350>.
- Marshall-Pescini, S., Whiten, A., 2008. Social learning of nut-cracking behavior in East African sanctuary-living chimpanzees (*Pan troglodytes schweinfurthii*). *J. Comp. Psychol.* 122 (2), 186–194. <https://doi.org/10.1037/0735-7036.122.2.186>.
- Martin, R.D., 1981. Relative brain size and basal metabolic rate in terrestrial vertebrates. *Nature* 293, 57–60. <https://doi.org/10.1038/293057a0>.
- Martin-Ordas, G., Haun, D., Colmenares, F., Call, J., 2010. Keeping track of time: evidence for episodic-like memory in great apes. *Anim. Cogn.* 13, 331–340.
- Matsuzawa, T., 1994. Field experiments on use of stone tools by chimpanzees in the wild. In: Wrangham, R.W., McGrew, W.C., de Waal, F.B.M., Heltne, P.G. (Eds.), *Chimpanzee Cultures*. Harvard University Press, Cambridge, pp. 351–370.
- Matsuzawa, T., 2009. Symbolic representation of number in chimpanzees. *Curr. Opin. Neurobiol.* 19, 1–7. <https://doi.org/10.1007/s10071-009-0282-4>.
- Matsuzawa, T., 2013. Evolution of the brain and social behavior in chimpanzees. *Curr. Opin. Neurobiol.* 23 (3), 443–449. <https://doi.org/10.1016/j.conb.2013.01.012>.
- Matsuzawa, T., Tomonaga, M., Tanaka, M., 2006. Cognitive Development in Chimpanzees. Springer-Verlag, Tokyo. <https://doi.org/10.1007/4-431-30248-4>.
- McEree, B., 2006. Accessing recent events. In: Ross, B.H. (Ed.), *Psychology of Learning and Motivation*. Elsevier Academic Press, Dordrecht, pp. 155–200. [https://doi.org/10.1016/S0079-7421\(06\)46005-9](https://doi.org/10.1016/S0079-7421(06)46005-9).
- McHenry, M.H., 2002. The fossil record of human ancestry. In: Hartwig, W.C. (Ed.), *The Primate Fossil Record*. Cambridge University Press, Cambridge, pp. 401–406. ISBN: 9780521663156.
- McPherron, S.P., Alemseged, Z., Marean, C.W., Wynn, J.G., Reed, D., Geraads, D., Bobe, R., Bérart, H.A., 2010. Evidence for stone-tool assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature* 466 (7308), 857–860. <https://doi.org/10.1038/nature09248>.
- Melis, A.P., Hare, B., Tomasello, M., 2006. Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Anim. Behav.* 72, 275–286. <https://doi.org/10.1016/j.anbehav.2005.09.018>.
- Melis, A.P., Hare, B., Tomasello, M., 2009. Chimpanzees coordinate in a negotiation game. *Evol. Hum. Behav.* 6, 381–392.
- Melis, A.P., Warneken, F., Jensen, K., Schneider, A., Call, J., Tomasello, M., 2011. Chimpanzees help conspecifics to obtain food and non-food items. *Proc. R. Soc. B: Biol. Sci.* 278, 1405–1413. <https://doi.org/10.1098/rspb.2010.1735>.
- Meltzoff, A.N., Brooks, R., 2008. Self-experience as a mechanism for learning about others: a training study in social cognition. *Dev. Psychol.* 44, 1257–1265. <https://doi.org/10.1037/a0012888>.
- Mendes, N., Hanus, D., Call, J., 2007. Raising the level: orangutans use water as a tool. *Biol. Lett.* 3, 453–455. <https://doi.org/10.1098/rsbl.2007.0198>.
- Mendes, N., Steinbeis, N., Bueno-Guerra, N., Call, J., Singer, T., 2018. Preschool children and chimpanzees incur costs to watch punishment of antisocial others. *Nat. Hum. Behav.* 2, 45–51. <https://doi.org/10.1038/s41562-017-0264-5>.
- Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63 (2), 81–97. <https://doi.org/10.1037/h0043158>.
- Miller, E.A., Buschman, T.J., 2015. Working memory capacity: limits on the bandwidth of cognition. *Daedalus* 112–122. [https://doi.org/10.1162/DAED\\_a.00320](https://doi.org/10.1162/DAED_a.00320).
- Miller, G.A., Galanter, E., Pribram, K.H., 1960. Plans and the Structure of Behavior. Holt, Rinehart, & Winston, New York. <https://doi.org/10.1037/10039-000>.
- Moher, M., Tuerk, A.S., Feigenson, L., 2012. Seven-month-old infants chunk items in memory. *J. Exp. Child Psychol.* 112, 361–377. <https://doi.org/10.1016/j.jecp.2012.03.007>.
- Moll, H., Tomasello, M., 2004. 12- and 18-month-old infants follow gaze to spaces behind barriers. *Dev. Sci.* 7, F1–F9. <https://doi.org/10.1111/j.1467-7687.2004.00315.x>.
- Montagu, M.F.A., 1955. Time, morphology, and neoteny in the evolution of man. *Am. Anthropol.* 57, 13–27. <https://doi.org/10.1525/aa.1955.57.1.02a00030>.
- Morey, C.C., Cowan, N., 2018. Can we distinguish three maintenance processes in working memory? *Ann. N. Y. Acad. Sci.* 1424, 45–51. <https://doi.org/10.1111/nyas.13925>.
- Mulcahy, N.J., Call, J., 2006. Apes save tools for future use. *Science* 312, 1038–1040. <https://doi.org/10.1126/science.1125456>.
- Mulcahy, N.J., Call, J., Dunbar, R.I.M., 2005. Gorillas (*Gorilla gorilla*) and orangutans (*Pongo pygmaeus*) encode relevant problem features in a tool-using task. *J. Comp. Psychol.* 119, 23–32. <https://doi.org/10.1037/0735-7036.119.1.23>.
- Nakashima, N., 2003. Developmental Changes of the Responses to the Vocalizations in Infant Chimpanzees. Unpublished Master Thesis. Primate Research Institute, Kyoto University in Japanese.
- Neadle, D., Bandini, E., Tennie, C., 2020. Testing the individual and social learning abilities of task-naïve captive chimpanzees (*Pan troglodytes sp.*) in a nut-cracking task. *PeerJ* 8, e8734. <https://doi.org/10.7717/peerj.8734>.
- Oberauer, K., 2013. The focus of attention in working memory — from metaphors to mechanisms. *Front. Hum. Neurosci.* 7, e00673 <https://doi.org/10.3389/fnhum.2013.00673>.
- Oberauer, K., 2018. Removal of irrelevant information from working memory: sometimes fast, sometimes slow, and sometimes not at all. *Ann. N. Y. Acad. Sci.* 1424, 239–255. <https://doi.org/10.1111/nyas.13603>.
- Okamoto-Barth, S., Call, J., Tomasello, M., 2007. Great apes’ understanding of other individuals’ line of sight. *Psychol. Sci.* 18, 462–468. <https://doi.org/10.1111/j.1467-9280.2007.01922.x>.
- Okamoto-Barth, S., Tomonaga, M., Tanaka, M., Matsuzawa, T., 2008. Development of using experimenter-given cues in infant chimpanzees: longitudinal changes in behavior and cognitive development. *Dev. Sci.* 11, 98–108. <https://doi.org/10.1111/j.1467-7687.2007.00658.x>.
- Olesen, P.J., Nagy, Z., Westerberg, H., Klingberg, T., 2003. Combined analysis of DTI and fMRI data reveals a joint maturation of white and grey matter in a fronto-parietal network. *Cogn. Brain Res.* 18, 48–57. <https://doi.org/10.1016/j.cogbrainres.2003.09.003>.
- Onishi, K.H., Baillargeon, R., 2005. Do 15-month-old infants understand false beliefs? *Science* 308 (5719), 255–258. <https://doi.org/10.1126/science.1107621>.
- Parker, S.T., McKinney, M.L., 1999. *The Evolution of Cognitive Development in Monkeys, Apes, and Humans*. The Johns Hopkins University Press, Baltimore.
- Passingham, R.E., Wise, S.P., 2012. *The neurobiology of the prefrontal cortex*. Anatomy, evolution, and the origin of insight. Oxford Psychology Series. Oxford University Press, Oxford.
- Peake, P.K., Hebl, M., Mischel, W., 2002. Strategic attention deployment for delay of gratification in working and waiting situations. *Dev. Psychol.* 38, 313–326. <https://doi.org/10.1037/0012-1649.38.2.313>.
- Perner, J., 1991. *Learning, Development, and Conceptual Change. Understanding the Representational Mind*. The MIT Press, Cambridge.
- Perner, J., Ruffman, T., 2005. Psychology, infants’ insight into the mind: How deep? *Science* 308 (5719), 214–216. <https://doi.org/10.1126/science.1111656>.
- Piaget, J., 1959. *The Language and Thought of the Child*, third ed. Routledge & Kegan Paul, London.
- Pilbeam, D., Gould, S.J., 1974. Size and scaling in human evolution: *homo sapiens* is a peculiar large primate; however, all australopithecines are versions of the “same” animal. *Science* 186 (4167), 892–901. <https://doi.org/10.1126/science.186.4167.892>.
- Postle, B.R., 2006. Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38. <https://doi.org/10.1016/j.neuroscience.2005.06.005>.
- Poti, P., Langer, J., Savage-Rumbaugh, S., Brakke, K.E., 1999. Spontaneous logicomathematical constructions by chimpanzees (*Pan troglodytes*, *P. paniscus*). *Anim. Cogn.* 2, 147–156. <https://doi.org/10.1007/s100710050035>.
- Potter, M.C., 2012. Recognition and memory for briefly presented scenes. *Front. Psychol.* 3, 32. <https://doi.org/10.3389/fpsyg.2012.00032>.
- Potter, M.C., Staub, A., O’Connor, D.H., 2004. Pictorial and conceptual representation of glimpsed pictures. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 478–489. <https://doi.org/10.1037/0096-1523.30.3.478>.
- Preuss, T.M., 2004. What is it like to be a human? In: Gazzaniga, M.S. (Ed.), *The Cognitive Neurosciences III*, 3rd ed. MIT Press, Cambridge, pp. 5–22.
- Read, D.W., 2006. Working memory: a cognitive limit to non-human primate recursive thinking prior to hominid evolution?. In: *Cognitive Science 2006. Proceedings. The 28th Annual Conference of the Cognitive Science Society*. July 26–29, 2006, Sheraton Vancouver Wall Center, Vancouver, British Columbia, Canada. Mahwah, NJ, Lawrence Erlbaum Associates, for the Cognitive Science Society, pp. 2674–2679. <https://doi.org/10.1177/147470490800600413>.
- Read, D.W., 2008. Working memory: a cognitive limit to non-human primate recursive thinking prior to hominid evolution. *Evol. Psychol.* 6, 676–714. <https://doi.org/10.1177/147470490800600413>.
- Read, D.W., 2010. From experiential-based to relational-based forms of social organization: a major transition in the evolution of *Homo sapiens*. In: Dunbar, R., Gamble, C., Gowlett, J. (Eds.), *Social Brain, Distributed Mind*. Oxford University Press, Oxford, pp. 1–15. <https://doi.org/10.5871/bacad/9780197264522.003.0010>.
- Read, D.W., 2012a. Cultural kinship as a computational system: from bottom-up to top-down forms of social organization. *Comput. Math. Organ. Theory* 18, 232–253. <https://doi.org/10.1007/s10588-012-9113-2>.
- Read, D.W., 2012b. *How Culture Makes Us Human: Primate Social Evolution and Formation of Human Societies*. Left Coast Press, Walnut Creek.
- Read, D.W., 2019. From past to present: the deep history of kinship. In: Saqalli, M., Vander Linden, M. (Eds.), *Integrating Qualitative and Social Science Factors in Archaeological Modelling*. Springer, Berlin, pp. 1–20. <https://doi.org/10.1007/978-3-030-12723-7>.
- Read, D.W., Andersson, C., 2019. Cultural complexity and complexity evolution. *Adapt. Behav.* 1–30. <https://doi.org/10.1177/1059712318822298>.
- Read, D.W., van der Leeuw, S., 2008. Biology is only part of the story. *Philosophical Transactions of the Royal Society Series B* 363, pp. 1959–1968. <https://doi.org/10.1098/rstb.2008.0002>.
- Read, D., van der Leeuw, S., 2015. The extension of social relations in time and space during the Palaeolithic and beyond. In: Wenban-Smith, F., Coward, F., Hosfield, R., Pope, M. (Eds.), *Settlement, Society, and Cognition in Human Evolution*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9781139208697.004>.
- Read, D.W., Lane, D.A., van der Leeuw, S., 2009. The innovation innovation. In: Lane, D.A., Pumain, D., van der Leeuw, S.E., West, G. (Eds.), *Complexity Perspectives in Innovation and Social Change*. Springer Verlag, Berlin, pp. 43–84. [https://doi.org/10.1007/978-1-4020-9663-1\\_3](https://doi.org/10.1007/978-1-4020-9663-1_3).
- Reis, A., Castro-Caldas, A., 1997. Illiteracy: a cause for biased cognitive development. *J. Int. Neuropsychol. Soc.* 3, 444–450.
- Reyes-Garcia, V., Pyhälä, A., Díaz-Reviriego, I., Duda, R., Fernández-Llamazares, A., Gallois, S., et al., 2016. Schooling, local knowledge and working memory: a study among three contemporary hunter-gatherer societies. *PLoS ONE* 11 (1), e0145265. <https://doi.org/10.1371/journal.pone.0145265>.
- Reznick, J.S., Morrow, J.D., Goldman, B.D., Snyder, J., 2004. The onset of working memory in infants. *Infancy* 6 (1), 145–154. [https://doi.org/10.1207/s15327078in0601\\_7](https://doi.org/10.1207/s15327078in0601_7).

- Rhodes, S., Cowan, N., 2018. Attention in working memory: attention is needed but it yearns to be free. *Ann. N. Y. Acad. Sci.* 1424, 52–63. <https://doi.org/10.1111/nyas.13652>.
- Rightmire, G.P., 2004. Brain size and encephalization in Early to Mid-Pleistocene *Homo*. *Am. J. Phys. Anthropol.* 124, 109–123. <https://doi.org/10.1002/ajpa.10346>.
- Rilling, J.K., 2006. Human and non-human primate brains: are they allometrically scaled versions of the same design? *Evol. Anthropol.* 15, 65–77. <https://doi.org/10.1002/evan.20095>.
- Rilling, J.K., Barks, S.K., Parr, L.A., Preuss, T.M., Faber, T.L., Pagnoni, G., Bremner, J.D., Votaw, J.R., 2007. A comparison of resting-state brain activity in humans and chimpanzees. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17146–17151. <https://doi.org/10.1073/pnas.0705132104>.
- Rilling, J.K., Scholz, J., Preuss, T.M., Glasser, M.F., Errangi, B.K., Behrens, T.E., 2012. Differences between chimpanzees and bonobos in neural systems supporting social cognition. *Soc. Cogn. Affect. Neurosci.* 7, 369–379. <https://doi.org/10.1093/scan/nsr017>.
- Rosati, A.G., Wobber, V., Hughes, K., Santos, L.R., 2014. Comparative developmental psychology: how is human cognitive development unique? *Evol. Psychol.* 12 (2), 448–473. <https://doi.org/10.1177/147470491401200211>.
- Ruff, C.B., Trinkhaus, E., Holliday, T.W., 1997. Body mass and encephalization in Pleistocene *Homo*. *Nature* 387, 173–176. <https://doi.org/10.1038/387173a0>.
- Russon, A.E., 2004. Great ape cognitive systems. In: Russon, A.E., Begun, D.R. (Eds.), *The Evolution of Thought: Evolutionary Origins of Great Ape Intelligence*. Cambridge University Press, Cambridge and New York, pp. 76–100. <https://doi.org/10.1017/CBO9780511542299.008>.
- Sakai, T., Mikami, A., Tomonaga, M., Matsui, M., Suzuki, J., Hamada, Y., Tanaka, M., Miyabe-Nishiwaki, T., Makishima, H., Nakatsukasa, M., Matsuzawa, T., 2012. Differential prefrontal white matter development in chimpanzees and humans. *Curr. Biol.* 21 (16), 1397–1402. <https://doi.org/10.1016/j.cub.2011.12.052>.
- Scarf, D., Ruffman, T., 2017. Great Apes' Insight into the Mind: How Great? [Online Commentary]. <https://science.sciencemag.org/content/354/6308/110/tab-e-letters>.
- Semaw, S., Renne, P., Harris, J.W.K., Feibel, C.S., Bernor, R.L., Fesseha, N., Mowbray, K., 1997. 2.5-Million-year-old stone tools from Gona, Ethiopia. *Nature* 385 (6614), 333–336. <https://doi.org/10.1038/385333a0>.
- Semaw, S., Rogers, M.J., Quade, J., Renne, P.R., Butler, R.F., Dominguez-Rodrigo, M., Stout, D., Hart, W.S., Pickering, T., Simpson, S.W., 2003. 2.6-million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *J. Hum. Evol.* 45 (2), 169–177. [https://doi.org/10.1016/s0047-2484\(03\)00093-9](https://doi.org/10.1016/s0047-2484(03)00093-9).
- Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., Van Hoesen, G.W., 2001. Prefrontal cortex in humans and apes: a comparative study of area 10. *Am. J. Phys. Anthropol.* 114, 224–241. [https://doi.org/10.1002/1096-8644\(200103\)114:3%3C224::AID-AJPA1022%3E3.0.CO;2-1](https://doi.org/10.1002/1096-8644(200103)114:3%3C224::AID-AJPA1022%3E3.0.CO;2-1).
- Shea, J.J., 2017. Occasional, obligatory, and habitual stone tool use in hominin evolution. *Evol. Anthropol.* 26 (5), 200–217. <https://doi.org/10.1002/evan.21547>.
- Siegel, L.S., Ryan, E.B., 1989. The development of working memory in normally achieving and subtypes of learning disabled children. *Child Dev.* 60, 973–980. <https://doi.org/10.2307/1131037>.
- Silberberg, A., Kearns, D., 2009. Memory for the order of briefly presented numerals in humans as a function of practice. *Anim. Cogn.* 12, 405–407. <https://doi.org/10.1007/s10071-008-0206-8>.
- Smaers, J.B., Gómez-Robles, A., Parks, A.N., Sherwood, C.C., 2017. Exceptional evolutionary expansion of prefrontal cortex in great apes and humans. *Curr. Biol.* 27 (5), 714–720. <https://doi.org/10.1016/j.cub.2017.05.015>.
- Somel, M., Liu, X., Tang, L., Yan, Z., Hu, H., Guo, S., Jiang, X., Zhang, X., Xu, G., Xie, G., et al., 2011. MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. *PLoS Biol.* 9, e1001214. <https://doi.org/10.1371/journal.pbio.1001214>.
- Southgate, V., Senju, A., Csibra, G., 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18, 587–592. <https://doi.org/10.1111/j.1467-9280.2007.01944.x>.
- Spear, L.P., 2000. The adolescent brain and age-related behavioral manifestations. *Neurosci. Biobehav. Rev.* 24, 417–463. [https://doi.org/10.1016/S0149-7634\(00\)00014-2](https://doi.org/10.1016/S0149-7634(00)00014-2).
- Spinozzi, G., Natale, F., Langer, J., Brakke, K.E., 1999. Spontaneous class grouping behavior by bonobos (*Pan paniscus*) and common chimpanzees (*Pan troglodytes*). *Anim. Cogn.* 2, 157–170. <https://doi.org/10.1007/s100710050036>.
- Stout, D., 2011. Stone tool making and the evolution of human culture and cognition. *Philos. Trans. R. Soc. Biol. Sci.* 366, 1050–1059. <https://doi.org/10.1098/rstb.2010.0369>.
- Surian, L., Geraci, A., 2012. Where will the triangle look for it? Attributing false beliefs to a geometric shape at 17 months. *Br. J. Dev. Psychol.* 30 (1), 30–44. <https://doi.org/10.1111/j.2044-835X.2011.02046.x>.
- Swanson, H.L., Beebe-Frankenberger, M., 2004. The relationship between working memory and mathematical problem solving in children at risk and not at risk for math disabilities. *J. Educ. Psychol.* 96, 471–491. <https://doi.org/10.1037/0022-0663.96.3.471>.
- Swanson, H.L., Sachse-Lee, C., 2001. Mathematical problem solving and working memory in children with learning disabilities: both executive and phonological processes are important. *J. Exp. Child Psychol.* 79, 294–321. <https://doi.org/10.1006/jecp.2000.2587>.
- Tecwyn, E.M., Thorpe, K.S., Chappel, J., 2013. A novel test of planning ability: great apes can plan step-by-step but not in advance of action. *Behav. Processes* 100, 174–184. <https://doi.org/10.1016/j.beproc.2013.09.016>.
- Tecwyn, E.M., Thorpe, K.S., Chappel, J., 2014. Development of planning in 4- to 10-year-old children: reducing inhibitory demands does not improve performance. *J. Exp. Child Psychol.* 125, 85–101.
- Tomasello, M., 1995. Joint attention as social cognition. In: Moore, C., Dunham, P.J. (Eds.), *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum, Hillsdale, pp. 103–130.
- Tomasello, M., 1999. *The Cultural Origins of Human Cognition*. Harvard University Press, London. <https://doi.org/10.2307/j.ctvjs4f4c>.
- Tomasello, M., 2014. *A Natural History of Human Thinking*. Harvard University Press, Cambridge.
- Tomasello, M., Carpenter, M., 2005. The emergence of social cognition in three young chimpanzees. *Monogr. Soc. Res. Child Dev.* 70, vii–132. <https://www.jstor.org/stable/3701393>.
- Toth, N., Schick, K., Semaw, S., 2006. A comparative study of the stone tool-making skills of *Pan*, *Australopithecus*, and *Homo sapiens*. In: Toth, N., Schick, K. (Eds.), *The Oldowan: Case Studies into the Earliest Stone Age*. Gosport, Indiana, Stone Age Institute, "Stone Age Institute Publication Series Number 1"; and Bloomington, Indiana, University of Indiana, pp. 156–222. ISBN 9780979227608.
- Trevarthen, C., 1979. Communication and cooperation in early infancy: a description of primary intersubjectivity. In: Bullowa, M. (Ed.), *Before Speech: The Beginning of Interpersonal Communication*. Cambridge University Press, Cambridge, pp. 321–348.
- Trick, L., Pylyshyn, Z., 1994. Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol. Rev.* 101, 80–102. <https://doi.org/10.1037/0033-295X.101.1.80>.
- Tsetlin, M.M., Novikova, S.I., Orekhova, E.V., Pushina, N.P., Malakhovskaya, E.V., Filatov, A.I., Stroganova, T.A., 2012. Developmental continuity in the capacity of working memory from infancy to preschool age. *Neurosci. Behav. Physiol.* 42 (7), 692–699. <https://doi.org/10.1007/s11055-012-9620-0>.
- Turing, A., 1936. On computable numbers, with an application to the Entscheidungsproblem. *Proc. London Math. Soc.* 42, 230–265. <https://doi.org/10.1112/plms/s2-42.1.230>.
- Ullman, H., 2019. Personal Communication. Email communication to D. Read.
- Ullman, H., Almeida, R., Klingberg, T., 2014. Structural maturation and brain activity predict future working memory capacity during childhood development. *J. Neurosci.* 34, 1592–1598. <https://doi.org/10.1523/JNEUROSCI.0842-13.2014>.
- van der Leeuw, S.E., 2020. *Social Sustainability, Past and Future: Undoing Unintended Consequences for the Earth's Survival*. Cambridge University Press, Cambridge.
- van der Leeuw, S.E., Lane, D.A., Read, D.W., 2009. The long-term evolution of social organization. In: Lane, D.A., Pumain, D., Van der Leeuw, S.E., West, G. (Eds.), *Complexity Perspectives in Innovation and Social Change*. Springer Verlag, Berlin, pp. 85–116. <https://doi.org/10.1007/978-1-4020-9663-14>.
- Villmoare, B., Kimbel, W.H., Seyoum, C., Campisano, C.J., DiMaggio, E., Rowan, J., Braun, D.R., Arrowsmith, J.R., Reed, K.E., 2015. Early *Homo* at 2.8 Ma from Ledi-Geraru, Afar, Ethiopia. *Science* 347 (6228), 1352–1355. <https://doi.org/10.1126/science.aaa1343>.
- Völter, C.J., Mundry, R., Call, J., Seed, A.M., 2019. Chimpanzees flexibly update working memory contents and show susceptibility to distraction in the self-ordered search task. *Proc. R. Soc. B: Biol. Sci.* 286 (1907), 20190715. <https://doi.org/10.1098/rspb.2019.0715>.
- Washburn, D.A., Gullledge, J.P., James, F., Rumbaugh, D.M., 2007. A species difference in Visuospace working memory: does language link what, with, where? *Int. J. Comp. Psychol.* 20, 55–64. Retrieved from: <https://escholarship.org/uc/item/5wf7b7fw>.
- Wellman, H.M., Cross, D., 2001. Theory of mind and conceptual change. *Child Dev.* 72, 702–707.
- Wimmer, H., Perner, J., 1983. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5).
- Wobber, V., Herrmann, E., Hare, B., Wrangham, R., Tomasello, M., 2014. Differences in the early cognitive development of children and great apes. *Dev. Psychobiol.* 56 (3), 547–573. <https://doi.org/10.1002/dev.21125>.
- Woodman, G.F., Luck, S.J., Schall, J.D., 2007. The role of working memory representations in the control of attention. *Cereb. Cortex* 17, 1118–1124.
- Woodward, A., 1998. Infants selectively encode the goal object of an actors reach. *Cognition* 69, 1–34.
- Wynn, T., Coolidge, F.L., 2004. The expert Neandertal mind. *J. Hum. Evol.* 46, 467–487. <https://doi.org/10.1016/j.jhevol.2004.01.005>.
- Zhou, X., Zhu, D., Qi, X.L., Li, S., King, S.G., Salinas, E., Stanford, T.R., Constantinidis, C., 2016. Neural correlates of working memory development in adolescent primates. *Nat. Commun.* 7, 13423. <https://doi.org/10.1038/ncomms13423>.